

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2005, 19:52:02 ; Search time 3007 Seconds

(Without alignments)
10393.620 Million cell updates/sec

Title: US-09-977-579-2-REV

Perfect score: 645

Sequence: 1 ATGCCGNCNTTAAVMGNNT.....SNGGNGTNCNGTNGARGAR 645

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapect 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_htg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pt: *
10: gb_ro: *
11: gb_scs: *
12: gb_sy: *
13: gb_un: *
14: gb_yt: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|--------------|--------------------|
| 1 | 466 | 72.2 | 1261 | 6 AR359850 | AR359850 Sequence |
| 2 | 466 | 72.2 | 1261 | 6 AX039100 | AX039100 Sequence |
| 3 | 466 | 72.2 | 1261 | 9 HS243396 | AU243396 Homo sapi |
| 4 | 466 | 72.2 | 5306 | 9 AB032984 | AB032984 Homo sapi |
| 5 | 462.8 | 71.8 | 645 | 6 AX048004 | AX048004 Sequence |
| 6 | 462.8 | 71.8 | 2220 | 6 AR359849 | AR359849 Sequence |
| 7 | 462.8 | 71.8 | 2220 | 6 AR359899 | AR359899 Sequence |
| 8 | 462.8 | 71.8 | 2220 | 10 RNO243395 | AJ243395 Rattus no |
| 9 | 462.8 | 71.8 | 2632 | 6 AX048005 | AX048005 Sequence |
| 10 | 462.8 | 71.8 | 3107 | 10 AF378093 | AF378093 Rattus no |
| 11 | 462.8 | 71.8 | 3108 | 6 AX047984 | AX047984 Sequence |
| 12 | 462.8 | 71.8 | 3296 | 6 AB097521 | AB097521 Macaca fa |
| 13 | 462.8 | 71.8 | 3910 | 10 BC070899 | BC070899 Rattus no |
| 14 | 461.2 | 71.5 | 4025 | 10 AK173115 | AK173115 Mus muscu |
| 15 | 461.2 | 71.5 | 4169 | 10 BC053919 | BC053919 Mus muscu |
| 16 | 461.2 | 71.5 | 4176 | 10 BC058636 | BC058636 Mus muscu |
| 17 | 460.6 | 71.4 | 670 | 10 AY049036 | AY049036 Mus muscu |
| 18 | 428.4 | 66.4 | 606 | 6 CQ728741 | CQ728741 Sequence |
| 19 | 367.2 | 56.9 | 922 | 5 BX935227 | BX935227 Gallus ga |

| | | | | | |
|----|-------|------|--------|--------------|--------------------|
| 20 | 323.8 | 50.2 | 912 | 5 BX931708 | BX931708 Gallus ga |
| 21 | 301 | 46.7 | 471 | 6 BD059018 | BD059018 Secreted |
| 22 | 300.4 | 46.6 | 1654 | 5 BC077295 | BC077295 Xenopus l |
| 23 | 165 | 25.6 | 127347 | 2 AP000682 | AP000682 Homo sapi |
| 24 | 165 | 25.6 | 144833 | 2 AC063921 | AC063921 Homo sapi |
| 25 | 165 | 25.6 | 149800 | 2 AC021981 | AC021981 Homo sapi |
| 26 | 165 | 25.6 | 172546 | 2 AC024604 | AC024604 Homo sapi |
| 27 | 165 | 25.6 | 174285 | 2 AC145484 | AC145484 Lemur cat |
| 28 | 165 | 25.6 | 178169 | 9 AP002765 | AP002765 Homo sapi |
| 29 | 165 | 25.6 | 178431 | 2 AP002749 | AP002749 Homo sapi |
| 30 | 165 | 25.6 | 181471 | 9 AC069539 | AC069539 Homo sapi |
| 31 | 162.6 | 25.2 | 140307 | 2 AC135353 | AC135353 Mus muscu |
| 32 | 162.6 | 25.2 | 238861 | 2 AC148331 | AC148331 Mus muscu |
| 33 | 162.4 | 25.2 | 214306 | 2 AC128723 | AC128723 Rattus no |
| 34 | 158.4 | 24.6 | 1335 | 9 HOMGSC1B | LI6242 Homo sapien |
| 35 | 158.4 | 24.6 | 1404 | 9 HOMGSC1B | LI0338 Human bodiu |
| 36 | 158.4 | 24.6 | 1525 | 9 BC067122 | BC067122 Homo sapi |
| 37 | 158 | 24.5 | 657 | 4 OC035382 | U35382 Oryctolagus |
| 38 | 158 | 24.5 | 657 | 6 AX048006 | AX048006 Sequence |
| 39 | 155.6 | 24.1 | 617 | 6 CQ722293 | CQ722293 Sequence |
| 40 | 152.2 | 23.6 | 1490 | 6 AX401978 | AX401978 Sequence |
| 41 | 152.2 | 23.6 | 1125 | 10 PATSCHB1B | M91808 Rattus norv |
| 42 | 151 | 23.4 | 1125 | 10 MM046681 | U46681 Mus musculu |
| 43 | 151 | 23.4 | 1340 | 10 MM085786 | U85786 Mus musculu |
| 44 | 151 | 23.4 | 1557 | 10 BC039140 | BC039140 Mus muscu |
| 45 | 151 | 23.4 | 1568 | 10 BC009652 | BC009652 Mus muscu |

ALIGNMENTS

RESULT 1
AR359850
LOCUS AR359850 Sequence 4 from patent US 6593565.
DEFINITION AR359850
ACCESSION AR359850
VERSION AR359850.1 GI:3766660
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1261)
AUTHORS Healin,P. and Lynnam,N.R.
TITLE Vehicle interior rearview mirror assembly including an
accessory-containing housing
JOURNAL Patent: US 6593565-A 4 15-JUL-2003;
FEATURES
source location/Qualifiers
1..1261
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 72.2%; Score 466; DB 6; Length 1261;
Best Local Similarity 58.8%; Pred. No. 1.2e-106;
Matches 379; Conservative 145; Mismatches 121; Indels 0; Gaps 0;
QY 1 ATGCCGNCNTTAAVMGNNTTCCNYTNGCNWSNTNGTNTATNTATYTGCTGWSN 60
DB 376 ATGCGCTCCCTCAATAGATTGTTCCCTCGGCTTCTCGGCTTATCTAGT 435
QY 61 GNTNGTTCGNGTNGTNGARGTNCNWSNBARCNARGCNGNRCGNAAYCN 120
DB 436 GTCTGCTCCCTGTGTGTGGAAGTCCCTCGAGAGCGAGCGGTGCGGCAACCC 495
QY 121 ATGAARYTMMGNTGYATNMTGATGAARMGNGARGGTNGARGCNACNMGNTGN 180
DB 496 ATGAAGTGGCGCTCATCTCTCGATGAGAGAGAGAGGTGAGCGCACACGCTGTG 555
QY 181 GARTGCTTAAVMGNCCNGARGNGNARGATTTTNTATNTAAGRTATMGAAVGN 240
DB 556 GAATGGTTCTAACGGCCGAGGGGCTTAAGATTCTTATTAAGAGTATCGGAAGGC 615
QY 241 CAYCARGGTNGARWNSCNTTTCARGMGNVNTNCARTGGAAYGNNMAARGAYTN 300

```

Db      616 CACACAGAGGTGAGAGCCCTTTACAGGGGCGCTGAGTGAATGACAGCAAGACCTG 675
Qy      301 CARGAGTGMNSAATNACNGTNTYNAAYGTNACNYTMAAYGAWNSNGANTYNTAYACNTGY 360
Db      676 CAGGAGGTGTCATCTGTGCTCAACGTCATCTGAACGACTCTGCGCTCTACACCTGC 735
Qy      361 AAYGTMSMNGARFTTYGARFTTYGARGCNCAYMNCNTTYGTNABACNACMNTN 420
Db      736 AATGTCTCCGGAGTTTGAAGGCGCATCGCCCTTTGTGAAGACGAGCGGCTG 795
Qy      421 ATNCCNYTMNGNTNACNGARGCNGNGARGAATYTYACMNSNGTNGTNSNGARATN 480
Db      796 ATCCCTTAAGATCACCGAGAGGCTGAGAGACTTCACTCTGTGTCTCAGAAATC 855
Qy      481 ATGATGTAAATYNTYNTGNTTYTNTYTNACNYTNGTYTNTATNABATGATNTATYTG 540
Db      856 ATGATGTACATCTCTGTGCTTCTTCACTCACTGTGTGCTCATGAGATGATATATTCG 915
Qy      541 TAYMGNARGTMSNABRGNGARGCNGNCARBARAAYGCNMSNGATYATYNTGCM 600
Db      916 TACAGAAAGTCTCAAAAAGCCGAGAGGACGCCCAAGAAAACCGTCTGACTTCTGCC 975
Qy      601 ATNCCMNSNGARAAYARBARAAYMSNGCNTNCGTNGTNGARGAR 645
Db      976 ATCCCATCTGAGAACAGAGAACTCTGCGGTACCACTGAGAGGA 1020

```

RESULT 2
 AX039100 1261 bp DNA linear PAT 18-NOV-2000
 LOCUS Sequence 4 from Patent WO0063367.
 DEFINITION AX039100
 ACCESSION AX039100
 VERSION AX039100.1 GI:11229276
 KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
 Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
 Cox, P., Dixon, A., Jackson, A. and Morgan, K.
 A novel family of beta sub-unit proteins from a voltage-gated sodi
 um channel, nucleic acids encoding them and therapeutic or.
 diagnostic uses there of
 Patent: WO 0063367-A 4 26-OCT-2000;

JOURNAL WARNER-LAMBERT COMPANY (US) ; Cambridge University Technical

FEATURES
 source Location/Qualifiers

1..1261
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Query Match 72.2%; Score 466; DB 6; Length 1261;
 Best Local Similarity 58.8%; Pred. No. 1.2e-106;
 Matches 379; Conservative 145; Mismatches 121; Indels 0; Gaps 0;

```

Qy      1 ATGCCGCGNTTAAAGNTNTTYCCNTYNGCMWSYNTGNTNTATNTATYTGCGTMSN 60
Db      376 ATGCCGCGCTTCAATGATGATGTTTCCCTGCTTCTCTGTCTTATCTAAGGTCAGT 435
Qy      61 GNTGTATYTCNGTNTGYGTNGARGTNCNWSNGARACNGARCGTNCARBGAAYCCN 120
Db      436 GTCTGCTTCCCTGTGTGTGTGAGAGTCCCTCGAAGACGAGGCGTGCAGGCAACCCC 495
Qy      121 ATGAARYTMMGTGATYTNMSNTGYATGAARMNGARGARCTNGARCNACNACNGTNTN 180
Db      496 ATGAAGCTGGCTGATCTCTCTGATGAAGAGAGAGAGGTCGAGGCCACACGAGTGTG 555
Qy      181 GATGTGTTATYTNMNCNNGARCGNGAARAGATTTTATNTATNTATYARTYMAAYGGA 240
Db      556 GAATGTGTTTACAGGCCCGAGGCGGTAAAGATTTCTTATTTACAGATATCGAATYGC 615

```

```

Qy      241 CAYCARGGTNGARBSNCCNTTYTCARGMGNTNCAATGTGAAYGMSNABARAYTN 300
Db      616 CACACAGAGGTGAGAGCCCTTTACAGGGGCGCTGCACTGGAATGACAGCAAGACCTG 675
Qy      301 CARGAGTGMNSAATNACNGTNTYNAAYGTNACNYTMAAYGAWNSNGANTYNTAYACNTGY 360
Db      676 CAGGAGGTGTCATCTGTGCTCAACGTCATCTGAACGACTCTGCGCTCTACACCTGC 735
Qy      361 AAYGTMSMNGARFTTYGARFTTYGARGCNCAYMNCNTTYGTNABACNACMNTN 420
Db      736 AATGTCTCCGGAGTTTGAAGGCGCATCGGCCCTTTGTGAAGACGAGCGGCTG 795
Qy      421 ATNCCNYTMNGNTNACNGARGCNGNGARGAATYTYACMNSNGTNGTNSNGARATN 480
Db      796 ATCCCTTAAGATCACCGAGAGGCTGAGAGACTTCACTCTGTGTCTCAGAAATC 855
Qy      481 ATGATGTAAATYNTYNTGNTTYTNTYTNACNYTNGTYTNTATNABATGATNTATYTG 540
Db      856 ATGATGTACATCTCTGTGCTTCTTCACTCACTGTGTGCTCATGAGATGATATATTCG 915
Qy      541 TAYMGNARGTMSNABRGNGARGCNGNCARBARAAYGCNMSNGATYATYNTGCM 600
Db      916 TACAGAAAGTCTCAAAAAGCCGAGAGGACGCCCAAGAAAACCGTCTGACTTCTGCC 975
Qy      601 ATNCCMNSNGARAAYARBARAAYMSNGCNTNCGTNGTNGARGAR 645
Db      976 ATCCCATCTGAGAACAGAGAACTCTGCGGTACCACTGAGAGGA 1020

```

RESULT 3
 HSA243396 1261 bp mRNA linear PRI 13-MAR-2000
 LOCUS HSA243396
 DEFINITION Homo sapiens mRNA for voltage-gated sodium channel beta-3 subunit
 (scn3b gene).
 ACCESSION AJ243396
 VERSION AJ243396.2 GI:7242612
 KEYWORDS scn3b gene; voltage-gated sodium channel beta-3 subunit.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
 Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
 Morgan, K., Stevens, E.B., Shaw, B., Cox, P., Dixon, A.K., Lee, K.,
 Pinnoch, R.D., Hughes, J., Richardson, P.J., Mizuguchi, K. and
 Jackson, A.P.
 beta 3: an additional auxiliary subunit of the voltage-sensitive
 sodium channel that modulates channel gating with distinct kinetics
 Proc. Natl. Acad. Sci. U.S.A. 97 (5), 2308-2313 (2000)

JOURNAL 20160948
 PUBMED 1068874

REFERENCE 2
 Morgan, K.
 Direct Submision
 Submitted (13-MAR-2000) Morgan K., Biochemistry, University of
 Cambridge, Tennis Court Road, Cambridge, CB2 1QW, UNITED KINGDOM
 Revised by [41]
 3 (bases 1 to 1261)

REMARK Morgan, K.
 Direct Submision
 Submitted (13-MAR-2000) Morgan K., Biochemistry, University of
 Cambridge, Tennis Court Road, Cambridge, CB2 1QW, UNITED KINGDOM
 On Mar 14, 2000 this sequence version replaced gi:7160974.

FEATURES
 source Location/Qualifiers
 1..1261
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="Brain"
 1..1261
 /gene="scn3b"
 <1..375
 /gene="scn3b"

gene
 5' UTR
 1..1261
 /gene="scn3b"

CDS

376..1024
/gene="scn3b"
/function="modulates channel gating kinetics"
/codon_start=1
/evidence=experimental
/product="voltage-gated sodium channel beta-3 subunit"
/protein_id="CB76825.1"
/db_xref="GI:7160975"
/db_xref="GDB:09N72"
/db_xref="UniProt/Swiss-Prot:Q9NY72"
translation="MPAFNRLEPPLASLVIWVSFCFVEBSEBAVQGNPKLKE
CISCKREBEATYWEWFPRDESGKDLIYEHRDGVSPFGRLQNNSSKLOLO
VSIITLVNTLNDGLYTCTNSRSEPEFAHRPVKTRILPIRVTEAGEDPTSVSEE
MMVITLVLTLMLEMLILEMYCYRKYSKAEEAAQENASDIALIPSENKENSAPVEE"
1025..>1261
/gene="scn3b"

3' UTR

ORIGIN

Query Match 72.2%; Score 466; DB 9; Length 1261;
Best Local Similarity 58.8%; Pred. No. 1.2e-106;
Matches 379; Conservative 145; Mismatches 121; Indels 0; Gaps 0;

Oy 1 ATGCCNGCNTTAAVMGNVTNTTTCCNTGNCNMSNGTYNTATNTAATYTGSTNMSN 60
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
Db 376 ATGCCTGCCTTCARAGATGTTCCTCCCTGGCTCTCTCGTCTATTACTACGGGTCAGT 435
|||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
Oy 61 GTNTGYTTYCCNGINTGYGTNGARGTNCNMSGNARACNGARGCNGTNCARGANAAYCCN 120
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
Db 436 GTCTGCTTCCCTGTGTGTGAAGTGCCCTCGAGACGAGAGCCGTGCAGAGCAACC 495
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
Oy 121 ATGARATNMNGTGTATMWSNTGYATGAARMNGARGRGTNGARAGCNACNACTNTGNTN 180
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
Db 496 ATGAAGCTGCGCTGTGATCTCTCCGACATGAAGAAGAGAGGTGAGGCCACCACTGTGTG 555
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
Oy 181 GARTGTTYYTAMGRCNGARGGNGNARGAATYYTNTATNTAATYGARTAVYMNAAYGNN 240
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
Db 556 GAATGTTCTACAGGCCCGGAGGCGGTAAAGATTTCTTATTACAGATATCGGAATGCG 615
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
Oy 241 CAYCARAGRTNGARMSNCNTTYCARGGNNMTYNTCARGTGAAYGAWMSNAARGAYYTN 300
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
Db 616 CACGAGAGGTGAGAGAGCCCCCTTTCAGGGGCGCTGCAGTGGAAATGGCAGCAAGGACCTG 675
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
Oy 301 CARGAATGMWSATTAACNGTNTTAAAYGTMACNTTMAAYGATWSNGSYNTTAYACTNGY 360
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
Db 676 CAGGACGTCATCATCTGCTCAACGTCATCTGAAACGACTCTGGCGCTCACACTGC 735
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
Oy 361 AAYGNMNGMGARTTGGARTTYGARONGCAVMGNCNTTYGTNAABACNAANNMTYN 420
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
Db 736 AAATGTCTCCGGAGATTTGAGTTTAGGGCACTGGCCCTTGTGAAGACGACGGCGCTG 795
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
Oy 421 ATNCCNTNMNGTNAACNGARGARGCNGNGARGAYTTYACNWSNGTNGTWSNGARATN 480
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
Db 796 ATCCCCCTTAAGGTACCGAGAGGCTGGAAGAGCACTCACTCTGTGTCTCAAAATC 855
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
Oy 481 ATGATGTATAINYNTNGTNTTYTTAACNTYNTGGYNTYNTAINAGARPATGATTAATYGY 540
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
Db 856 ATGAATGATCATCTCTGTCTTCTCTCAACCTGTGCTGCTCATGAGATGATATATGTC 915
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
Oy 541 TAYMNARAGTNWSNAARGCNGARGARGCNGCNGCARGAAPAAAGCWNSNGATYAAYTNCN 600
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
Db 916 TACAAAAAGGTCTCAAAGCGGAAGGAGGCCAAGAAAACGCTCTGACTACCTTGCC 975
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
Oy 601 ATNCCNWSNGAAPAAVARGARAAVWSNGCNGTNCNGTNGARGAR 645
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
Db 976 ATCCCATCTGAACAACAGAGAACTCTGGGTACCAAGTAGAGGAA 1020
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::

RESULT 4
AB032984 5306 bp mRNA linear PRI 11-NOV-1998
LOCUS Homo sapiens mRNA for KIAA1158 protein, partial cds.
DEFINITION
ACCESSION AB032984
VERSION AB032984.1 GI:6330135

KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS
Hirasawa,M., Nagase,T., Ishikawa,K., Kikuno,R., Nomura,N. and
Ohara,O.
TITLE
Characterization of cDNA clones selected by the Genemark analysis
from size-fractionated cDNA libraries from human brain
JOURNAL
DNA Res. 6 (5), 329-336 (1999)
MEDLINE
20039618
PubMed
10574461
REFERENCE
2 (bases 1 to 5306)
AUTHORS
Ohara,O., Nagase,T. and Kikuno,R.
TITLE
Direct Submission
JOURNAL
Submitted (04-OCT-1999) Otsu Ohara, Kazusa DNA Research Institute,
Laboratory of DNA Technology; 153-3 Yata, Kisarazu, Chiba
292-0812, Japan (E-mail:cdna1nf@kazusa.or.jp,
URL:http://www.kazusa.or.jp/huge/, Tel:+81-438-52-3913,
Fax:+81-438-52-3914)
FEATURES
source
location/Qualifiers
1..5306
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="hj00081"
/sex="male"
/issue_type="brain"
/clone_id="PBLuescriptII SK plus"
/dev_stage="adult"
1..5306
/gene="KIAA1158"
<1..695
/gene="KIAA1158"
/note="Start codon is not identified."
/codon_start=3
/product="KIAA1158 protein"
/protein_id="BAA6472.1"
/db_xref="GI:6330136"
/translation="RWSGSLGSGRQPKPAFNRLLPLASLVLIYWSYCFPVCEV
PSEHFAVQGNPMKLRCTSCMKREVEATTVEMPRGGDLIYVRNGHOVESP
FQRLQWNGSKDLQDVSYITLVPLTNDLSGLYCNVSEPEFPAARPPVYKTRLLPLRV
TEEGEDFTSVSEIMYITLVPLTLLIEMITCYRKVSKAEBAQENMSDYLAIPS
ENKENSAPVEE"
ORIGIN
Query Match 72.2%; Score 466; DB 9; Length 5306;
Beeb Local Similarity 58.8%; Pred. No. 1,86-106;
Matches 379; Conservative 145; Mismatches 121; Indels 0; Gaps 0;
QY 1 ATGCCGCGTAAATYAAVMGNYTNTTCCNYTNGCMNSNTYNTATNTAYTGGTWSN 60
DBB 48 ATGCTGCGCTTCAATAGATTGTTTCCCTCGCTTCTCTCGTCTATTCTAGTGGTCA 107
QY 61 GTNTGYTTTCCGNTGTGTGTGATGATGTCNCWNSGARAACNGAGCGTNGCARGNAAYCCN 120
DB 108 GTCTGCTTCCCTGTGTGTGTGGAAGTCCCTCGAGACGAGAGCGCGTGCAGGGCAACCCC 167
QY 121 ATGAATTTMGNTGTATNTMSYTGATATGAARAGNARBARAGTNGARGCNAACNAGTNGTN 180
DB 168 ATGAAGCTGCGCTCACTCTCTGTGATGAAGAGAGAGAGAGTGGAGGCCACCAAGTGGTG 227
QY 181 GATGAGTTTAAATMNCNGCNGARGGAGNAARGAATTTTATNTATNTAYGARTAYMNAAYG 240
DB 228 GAATGCTTCTACAGAGCCCGAGGCGGCTTAAGATTTCTTATTTCAGATATCGAATATGCG 287
QY 241 CAYCARGAGTNGARMSNCCNTTTCARGGMMGNTYNTCARTGAAVGGNWSNAARGAYTN 300
DB 288 CACCAAGAGGATGAGAGAGCCCTTTCAGGGGGCGCTGACATGGAATGCAAGAACGCTG 347
QY 301 CARGATGMSNATNACNGTNTTAAATGTAACNTTAAATGAAATYMSNGNTYNTATACNTGY 360

Db 348 CAGAGGTGTCATCACTGTGCTCAAGTCACTGTAAGACTCTGGCCCTACACCTGC 407
QY 361 AAYGTWMSNGNGARTTYGARITTYGARGCNCAVMGNCNTTYGTNABACNA CMNGNTN 420
Db 408 AATGTGTCAGGAGTTCAGAAATTCAGAGCAACAGGCTTTTGTGAAGCACAGACTG 420
QY 421 ATNCNVTMMGNTGNCNARGARGCGNGNGARGATYTTACNWSNCTNGTNSNGARATN 480
Db 468 ATCCCCCTTAAGATCACCGAGAGGCTGAGAGAGACTTCACTCTGTGTGCTCAGAAATC 527
QY 481 ATGATGTAATNTYNTNGTNTTYTACNVTNTGTYTNTAATNGARATGATNTATYTG 540
Db 528 ATGATGTAATCTCTGCTGTCTCTCTCACTGCTGCTCTCACTGAGATGATATATGTC 587
QY 541 TAYMGNABGTWMSNABRGNGARGCGNGNCARGABAAVGCNMSNGATYATYNTGCN 600
Db 588 TACAGAAAGTCTCAAAAGCCGAGAGGACGCCAAGAAAACGGTCTGACTACCTTGC 647
QY 601 ATNCNWSNGARAAVARGARAAVMSNGCNTNCNGTNGARGAR 645
Db 648 ATCCCACTGAGAACAGAGAACTCTGCGGTACCACTGAGAGAA 692

RESULT 5
AX048004 645 bp DNA linear PAT 15-DEC-2000
LOCUS Sequence 21 from Patent WO0069912.
DEFINITION AX048004
ACCESSION AX048004
VERSION AX048004.1 GI:11876882
KEYWORDS
SOURCE
ORGANISM
Lepus sp.
Lepus sp.
Bukayofa; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Lagomorpha; Leporidae; Lepus.
REFERENCE
AUTHORS
TITLE
JOURNAL
Curtis, R.A.
Gene encoding a sodium channel beta-3 subunit protein
Patent: WO 0069912-A 21 23-NOV-2000;
Millennium Pharmaceutical, Inc. (US)
FEATURES
source
1. 645
/organism="Lepus sp."
/mol_type="unassigned DNA"
/db_xref="taxon:137772"

ORIGIN

Query Match 71.8%; Score 462.8; DB 6; Length 645;
Best Local Similarity 58.4%; Pred. No. 6,7e-106;
Matches 377; Conservative 145; Mismatches 123; Indels 0; Gaps 0;

QY 1 ATGCCGNCNTTYAAVMGNTNTTYCCNYTNGCNMSNYTNGTNTAATNTATYTGGSN 60
Db 1 ATGCTGCTTCAACAGATTCCTCCCTAGCTTCTTAAGTCTCACTACCTGAGTCA 60
QY 61 GINTGYTTYCCNGTNTGTYTNGARGTNCNWSNGARACNGARGCNGTNCARGNAAYCN 120
Db 61 GTCGCTTCCCTGTGTGTGAAGTCCCTCGAGACAGAAAGGGTGCAGGCAATCCC 120
QY 121 ATGAARTTNGNTGYATNTSNTGYATGAARMGNGARGARGTNGARGCNAACNGTNGT 180
Db 121 ATGAAGCTGAGTGATCTCTGCTGATGAAGAGGAGGATGAGGCACTGTGTG 180
QY 181 GARTGTTTAYMGNCCNARGGNGNARGATYTTNTAATNTAYGARTAYMGNAAVGN 240
Db 181 GAGTGTTCATCAGGCTGAGGGCGGTAAAGATTTCTTATATATAGATTCGGAATG 240
QY 241 CAYCARGGTNGARMSNCNTTYCARGNMGYNTCARTGAAVGSNMAARGAYTN 300
Db 241 CACCAAGAGTGAAGAGCCCTTCAGAGCGCTCTGCACTGAGTGAATGAGCAAAAGCTG 300
QY 301 CARGAYGTWMSNATNACNGTNTAAYGTNACNTTAAVAYMSNGNTNTATYACNTGY 360
Db 301 CAGGAGCTATCATCACTGATCAATGATCACTTGAATGATCTGCGCTCTACACATGC 360

QY 361 AAYGTWMSNGNGARTTYGARITTYGARGCNCAVMGNCNTTYGTNABACNA CMNGNTN 420
Db 361 AATGTGTCAGGAGTTCAGAAATTCAGAGCAACAGGCTTTTGTGAAGCACAGACTG 420
QY 421 ATNCNVTMMGNTGNCNARGARGCGNGNGARGATYTTACNWSNCTNGTNSNGARATN 480
Db 421 ATACCTTGAGATCACTGAAGGCGGAGAAAGCTTCACTCCGTGTCTCGAAATC 480
QY 481 ATGATGTAATNTYNTNGTNTTYTACNVTNTGTYTNTAATNGARATGATNTATYTG 540
Db 481 ATGATGTAATCTCTGCTGTCTCTCTCACTGCTGCTCTCACTGAGATGATATATGTC 540
QY 541 TAYMGNABGTWMSNABRGNGARGCGNGNCARGABAAVGCNMSNGATYATYNTGCN 600
Db 541 TACAGAAAGTCTCAAAAGCCGAGAGGACGCCAAGAAAACGGTCTGACTACCTTGC 600
QY 601 ATNCNWSNGARAAVARGARAAVMSNGCNTNCNGTNGARGAR 645
Db 601 ATCCCTTCAAGAAACAGAGAACTCTGTGTACTGTGAGAGAA 645

RESULT 6
AR359849 2220 bp DNA linear PAT 17-AUG-2003
LOCUS Sequence 3 from patent US 6593565.
DEFINITION AR359849
ACCESSION AR359849
VERSION AR359849.1 GI:33766659
KEYWORDS
SOURCE
ORGANISM
Unknown.
Unclassified.
REFERENCE
AUTHORS
TITLE
JOURNAL
Hehlin, P. and Lymam, N.R.
Vehicle interior rearview mirror assembly including an
accessory-containing housing
Patent: US 6593565-A 3 15-JUL-2003;
FEATURES
source
1. 2220
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 71.8%; Score 462.8; DB 6; Length 2220;
Best Local Similarity 58.4%; Pred. No. 9,3e-106;
Matches 377; Conservative 145; Mismatches 123; Indels 0; Gaps 0;

QY 1 ATGCCGNCNTTYAAVMGNTNTTYCCNYTNGCNMSNYTNGTNTAATNTATYTGGSN 60
Db 363 ATGCTGCTTCAACAGATTCCTCCCTAGCTTCTTAAGTCTCACTACCTGAGTCA 422
QY 61 GINTGYTTYCCNGTNTGTYTNGARGTNCNWSNGARACNGARGCNGTNCARGNAAYCN 120
Db 423 GTCGCTTCCCTGTGTGTGAAGTCCCTCGAGACAGAAAGGGTGCAGGCAATCCC 482
QY 121 ATGAARTTNGNTGYATNTSNTGYATGAARMGNGARGARGTNGARGCNAACNGTNGT 180
Db 483 ATGAAGCTGAGTGATCTCTGCTGATGAAGAGGAGGAGTGAAGCCCACTGTGTG 542
QY 181 GARTGTTTAYMGNCCNARGGNGNARGATYTTNTAATNTAYGARTAYMGNAAVGN 240
Db 181 GAGTGTTCATCAGGCTGAGGGCGGTAAAGATTTCTTATATATAGATTCGGAATG 602
QY 241 CAYCARGGTNGARMSNCNTTYCARGNMGYNTCARTGAAVGSNMAARGAYTN 300
Db 603 CACCAAGAGTGAAGAGCCCTTCAGAGCGCTCTGCACTGAGTGAATGAGCAAAAGCTG 662
QY 301 CARGAYGTWMSNATNACNGTNTAAYGTNACNTTAAVAYMSNGNTNTATYACNTGY 360
Db 663 CAGGAGCTATCATCACTGATCAATGATCACTTGAATGATCTGCGCTCTACACATGC 722
QY 361 AAYGTWMSNGNGARTTYGARITTYGARGCNCAVMGNCNTTYGTNABACNA CMNGNTN 420

| | | | | |
|--|--|---|--|------|
| Oy | | 361 | AAVETNWSNMNGARRTTYGARCTTGGARGCNCANVMGNCNTTTYGTNAABACNAGMGNVTN | 420 |
| Dd | | 723 | AATGTTCACAGGGAGTTTGAAATTGCAGGACACAGGCCCTTTTGGAAGACCACAGACTG | 782 |
| Oy | | 421 | ATTNCCNYTNMGNTNACNGARGCGCNGCNGARGAYTTTACMWSNGTNGTWMSNGARATN | 480 |
| Dd | | 783 | ATACTCTTGCAAGTCACTGAAAGAAGCGGAGAAAGACTTCACCTCGTGCTCCGAAATC | 842 |
| Oy | | 481 | ATATGTATATNTYNTYNTNGTTTTYYTAACNTYTNGTNTMTAATNGARATATTAATGY | 540 |
| Dd | | 843 | ATGATGATACATCCTTCGTGCTTCTCTCACTTGTCGTTTATTTGATGATATCTAATTC | 902 |
| Oy | | 541 | TAYVGNARGTWNSNARGCNGARGCGCNGCNCARAPAAVAGCMWSNGATYAAYTNGCN | 600 |
| Dd | | 903 | TACGAAAAGCTCTTAAGCGCGGACGACGACAAGAAAAATGCTGTGATACCTTGCT | 962 |
| Oy | | 601 | ATTCCMWSNGBARAAYARBARAAVWSNGCNGTNCNGTNGTNGARGAR | 645 |
| Dd | | 963 | ATTCCTTCAGAGAACAGAGAACTCTGTGTACTCTGTGAGGAA | 1007 |
| RESULT 8 RNO243395 LOCUS DEFINITION | | RNO243395 | 2220 bp mRNA linear ROD 29-JAN-2003 | |
| ACCESSION | | AFJ243395 | | |
| VERSION | | AJ243395.2 | GI:7242802 | |
| KEYWORDS | | scn3b gene; voltage-gated sodium channel beta-3 subunit. | | |
| SOURCE | | Rattus norvegicus (Norway rat) | | |
| ORGANISM | | Rattus norvegicus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus. | | |
| REFERENCE | | 1 | Morgan, K., Stevens, E.B., Shaw, B., Cox, P., Dixon, A.K., Lee, K., Plimock, R.D., Hignes, V., Richardson, P.O., Mizuguchi, K. and Jackson, A.P. | |
| AUTHORS | | | beta 3 : an additional auxiliary subunit of the voltage-sensitive sodium channel that modulates channel gating with distinct kinetics Proc. Natl. Acad. Sci. U.S.A. 97 (5), 2308-2313 (2000) | |
| JOURNAL | | PUBMED | 10688874 | |
| REFERENCE | | 2 | Morgan, K. | |
| AUTHORS | | | Direct Submission | |
| TITLE | | | Submitted (28-JUN-1999) Morgan K., Biochemistry, University of Cambridge, Tennis Court Road, Cambridge, CB2 1QW, UNITED KINGDOM | |
| JOURNAL | | | Revised by [3] | |
| REMARK | | | 3 (bases 1 to 2220) | |
| REFERENCE | | | Morgan, K. | |
| AUTHORS | | | Direct Submission | |
| TITLE | | | Submitted (09-MAR-2000) Morgan K., Biochemistry, University of Cambridge, Tennis Court Road, Cambridge, CB2 1QW, UNITED KINGDOM | |
| JOURNAL | | | On Mar 14, 2000 this sequence version replaced gi:7161888. | |
| COMMENT | | | location/Qualifiers | |
| FEATURES | | | 1..2220 | |
| SOURCE | | | /organism="Rattus norvegicus" | |
| | | | /mol_type="mRNA" | |
| | | | /db_xref="taxon:10116" | |
| | | | /tissue_type="brain" | |
| | | | /tissue_1lb="lambda ZAP" | |
| | | | 1..2220 | |
| | | | /gene="scn3b" | |
| | | | <1..362 | |
| | | | /gene="scn3b" | |
| | | | 363..1010 | |
| | | | /gene="scn3b" | |
| | | | /function="modulates channel opening kinetics" | |
| | | | /codon_start=1 | |
| | | | /evidence=experimental | |
| | | | /product="voltage-gated sodium channel beta-3 subunit" | |
| gene | | | | |
| 5'UTR | | | | |
| CDS | | | | |

REFERENCE 1 (bases 1 to 3107)
AUTHORS Qu,Y., Curtis,R., Lawson,D., Gilbride,K., Ge,P., DiStefano P.S.,
TITLE Sli05-Santiago,I., Caterall,W.A. and Scheuer,T.
DIFFERENTIAL MODULATION OF SODIUM CHANNEL GATING AND PERSISTENT
SODIUM CURRENTS BY THE BETA1, BETA2, AND BETA3 SUBUNITS
Mol. Cell. Neurosci. 18 (5), 570-580 (2001)
JOURNAL 21919106
MEDLINE 11922146
REFERENCE 2 (bases 1 to 3107)
AUTHORS Curtis,R.
TITLE Direct Submission
JOURNAL Submitted (05-MAY-2001) Neurobiology, Millennium Pharmaceuticals
Inc., 75 Sidney Street, Cambridge, MA 02139, USA
FEATURES
SOURCE
1. 3107
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/tissue_type="dorsal root ganglion"
/dev_stage="adult"
1. 3107
/gene="SCN3B"
77..724
/gene="SCN3B"
/codon_start=1
/product="sodium channel beta 3 subunit"
/protein_id="AAK55415.1"
/db_xref="GI:14165176"
/translation="MPAPRNLPLASLVLYWVPCVPCVPSPTSEAVQGNPKLR
CISMKREVEATTVKFRPREGKDLITRYNGHOEVSPPQRILOMGSLOLD
VSIIVLVTLNDGLYTCNVSREFEPEHAFVPTTRILPLRLEAGEDETSVSEI
MMYILVFLTLMPLFEMIYCYRKVSKAEAAQENASDYLAIPSENKENVVPEE"

ORIGIN

Query Match 71.8%; Score 462.8; DB 10; Length 3107;
Best Local Similarity 58.4%; Pred. No. 1e-105;
Matches 377; Conservative 145; Mismatches 123; Indels 0; Gaps 0;
QY 1 ATGCCGNCNTTYAAYMGNYNTTTCNTYNGCNMSNTYNTNATNTATYATGGTWSN 60
DB 77 ATGCTGCTTCAACAAGATTCCTTCCCTAGCTTCTAGTGCATCTACTGCGTCA 136
QY 61 GTNTGTYTTCNGTNTGTYNGARNTNCNWSNGARACNGAGCGTNCARGNAAVCCN 120
DB 137 GTCTGCTTCCCTGTGTGTGGAAGTGCCCTCGAGACAGAAAGCGGTCAAGGCAATCCC 196
QY 121 ATGAARTNMGNTGYATNMSNTGYATGAABMGNGARAGRTNGARACNACNAGTNGTN 180
DB 197 ATGAAGCTGAGTGCACTCTCTGCATGAAGAGGAGGAGTGAGGCCACCACTGTGTG 256
QY 181 GARTGTTTAYMGNCCNGARNGNGNARAGAYTTTNTATNTATYARTAYMGNAAVGN 240
DB 257 GAGTGTCTTCAAGGCGGTGAAGGCGGTAAAGATTTCTTATATATGATATCGAATGCG 316
QY 241 CAYCARGARGTNGARWSNCCNTTYCARGNMGNNTNARTGAAVYGNWSNARAGAYTN 300
DB 317 CACCAAGAAAGTGAGAGCCCTTCCAAAGCGGTCTGACAGTGAATGGAGCAAAAGACTG 376
QY 301 CARAGYGTNMSNATNACNGNTYNTAAAGTNAANTNAYGAYMNGNGYNTTAAACNTGY 360
DB 377 CAGAGCGATTCATCACTGATCAATGTCATTTGAATGACTCTGGGCTCTTACACATGC 436
QY 361 AAYGTNMSNMGARTTYGARTTYGARGNCAYMGNCCNTTYGTAARAACNMGNYTN 420
DB 437 AATGTGTCAAGGAGGATTCGAATGAGGACACAGGCGCTTTTGTGAAGACACAGAGCTG 496
QY 421 ATNCCNTTNGNNTNACNGARAGCNGAGNGARAGAYTTTACNWSNGTNGTWSGARATN 480
DB 497 ATACCTTTCGAGCTCACTGAAGGCGGAGAAAGACTTCACTCCGTGTGTCTCGGAATTC 556
QY 481 ATGATGTAAYANTNNTYNTGNTTYYTNAACNTNTNGYNTNATNAGARATGATNTATGY 540

DB 557 ATGATGTAACCTTCCCTGCTTCTTCACTTGTGGCGTGTATGATGATCATATGCG 616
QY 541 TAYMGNARGTNMSNARAGCNGARNGCNGCARARAAVYGNWSNATAYTNGCN 600
DB 617 TACGAAAGGCTCTTAAGCGGAGAGGAGGACGACAGAAAATGCTGACTACTGCT 676
QY 601 ATNCCNMSNGARAAVYARAGAAAYMWSNCGTNCNGTNGARGAR 645
DB 677 ATCCCTTCAGAGAAACAAGAGAACTGTGTGACTGTGAGGAA 721

RESULT 11
AX047984 3108 bp DNA linear PAT 15-DEC-2000
LOCUS
DEFINITION Sequence 1 from Patent WO0069912.
AX047984
ACCESSION
VERSION AX047984.1 GI:11876881
KEYWORDS
SOURCE
ORGANISM
Rattus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1
AUTHORS Curtis,R.A.
TITLE Gene encoding a sodium channel beta-3 subunit protein
JOURNAL Patent: WO 0069912-A 1 23-NOV-2000;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
SOURCE
1. 3108
/organism="Rattus sp."
/mol_type="unassigned DNA"
/db_xref="taxon:10118"

ORIGIN

Query Match 71.8%; Score 462.8; DB 6; Length 3108;
Best Local Similarity 58.4%; Pred. No. 1e-105;
Matches 377; Conservative 145; Mismatches 123; Indels 0; Gaps 0;
QY 1 ATGCCGNCNTTYAAYMGNYNTTTCNTYNGCNMSNTYNTNATNTATYATGGTWSN 60
DB 78 ATGCTGCTTCAACAAGATTCCTTCCCTAGCTTCTAGTGCATCTACTGCGTCA 137
QY 61 GTNTGTYTTCNGTNTGTYNGARNTNCNWSNGARACNGAGCGTNCARGNAAVCCN 120
DB 138 GTCTGCTTCCCTGTGTGTGGAAGTGCCCTCGAGACAGAAAGCGGTCAAGGCAATCCC 197
QY 121 ATGAARTNMGNTGYATNMSNTGYATGAABMGNGARAGRTNGARACNACNAGTNGTN 180
DB 198 ATGAAGCTGAGTGCACTCTCTGCATGAAGAGGAGGAGTGAGGCCACCACTGTGTG 257
QY 181 GARTGTTTAYMGNCCNGARNGNGNARAGAYTTTNTATNTATYARTAYMGNAAVGN 240
DB 258 GAGTGTCTTCAAGGCGGTGAAGGCGGTAAAGATTTCTTATATATGATATCGAATGCG 317
QY 241 CAYCARGARGTNGARWSNCCNTTYCARGNMGNNTNARTGAAVYGNWSNARAGAYTN 300
DB 318 CACCAAGAAAGTGAGAGCCCTTCCAAAGCGGTCTGACAGTGAATGGAGCAAAAGACTG 377
QY 301 CARAGYGTNMSNATNACNGNTYNTAAAGTNAANTNAYGAYMNGNGYNTTAAACNTGY 360
DB 378 CAGAGCGATTCATCACTGATCAATGTCATTTGAATGACTCTGGGCTCTTACACATGC 437
QY 361 AAYGTNMSNMGARTTYGARTTYGARGNCAYMGNCCNTTYGTAARAACNMGNYTN 420
DB 438 AATGTGTCAAGGAGGATTCGAATGAGGACACAGGCGCTTTTGTGAAGACACAGAGCTG 497
QY 421 ATNCCNTTNGNNTNACNGARAGCNGAGNGARAGAYTTTACNWSNGTNGTWSGARATN 480
DB 498 ATACCTTTCGAGCTCACTGAAGGCGGAGAAAGACTTCACTCCGTGTGTCTCGGAATTC 557
QY 481 ATGATGTAAYANTNNTYNTGNTTYYTNAACNTNTNGYNTNATNAGARATGATNTATGY 540

| | | | |
|------------|--|---|-----------------------------|
| Oy | 541 | TATWGNAAAGTMSNARGCNGARGCNCNCAAGAPAAVGCNSNGATAYTNCGN | 600 |
| Db | 618 | TACAGAAAAGGCTCTTAAGCCGAAGGACACACAGGAAAAATGGCTCTGACTACTTGCT | 677 |
| Oy | 601 | ATTNCNMSGARAAVAARGAFAAVWSNGCNGTCNCNGTNGARGAR | 645 |
| Db | 678 | ATCCCTTCAGAGAACAGAGAAACTCTGTGTACTCTGTGAGGAA | 722 |
| RESULT 12 | | | |
| LOCUS | AB097521 | 3396 bp | mRNA linear PRI 06-DEC-2002 |
| DEFINITION | Macaca fascicularis brain cDNA clone:QmoA-13657, full insert sequence. | | |
| ACCESSION | AB097521 | | |
| VERSION | AB097521.1 | GI:26449236 | |
| KEYWORDS | oligo capping; fis (full insert sequence). | | |
| SOURCE | Macaca fascicularis (crab-eating macaque) | | |
| ORGANISM | Macaca fascicularis Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae; Macaca. | | |
| REFERENCE | 1 | Osada,N., Hida,M., Kusuda,J., Tanuma,R., Iseki,K., Hirata,M., Suto,Y., Hirai,M., Teroo,K., Suzuki,Y., Sugano,S. and Hashimoto,K. Assignment of 118 novel cDNAs of cynomolgus monkey brain to human chromosomes | |
| AUTHORS | Gene 275 (1), 31-37 (2001) | | |
| TITLE | 21458551 | | |
| JOURNAL | 11574149 | | |
| MEDLINE | 2 (bases 1 to 3296) | | |
| PUBMED | Hashimoto,K., Osada,N., Hida,M., Kusuda,J. and Sugano,S. Direct Submission | | |
| REFERENCE | Submitted (05-DEC-2002) Katsuyuki Hashimoto, National Institute of Infectious Diseases, Division of Genetic Resources; 23-1, Toyama 1-chome, Shinjuku-ku, Tokyo 162-8640, Japan (E-mail:Khashi@nih.go.jp, URL:http://www.nih.go.jp/yoken/genedbank/, Tel:-81-3-5285-1111(ex.2120), Fax:81-3-5285-1181) | | |
| AUTHORS | Lab host: TOPIO | | |
| TITLE | Vector: pMT8S-FL3 (Acc.No. AB009864) | | |
| JOURNAL | R. Site1: DraIII (CACTGCTGTC) R. Site2: DraIII (CACCATGTG) Description: 1st strand cDNA was primed with an oligo(dT) primer [ATGTCGCCCTTTTCTTTTTTTTT]; double-stranded cDNA was synthesized using specific 5' and 3' primers and amplified by PCR. The PCR product was digested with SfiI and size selection was performed to exclude fragments <1.5kb.The SfiI-digested PCR product was cloned into distinct DraIII-sites of pMT8S-FL3. XhoI sites just outside the DraII sites can be used to isolate the cDNA insert. Libraries were constructed by oligo-capping method. Custom primers used for sequencing (5' end primer [CTTTCGCTCTTAAAGCTGCG] ; 3' end primer [CGACCTGAGCTCGACGACA]). Location/Qualifiers | | |
| COMMENT | 1..3296 | | |
| FEATURES | /organism="Macaca fascicularis" | | |
| SOURCE | /mol_type="mRNA" | | |
| | /db_xref="taxon:9541" | | |
| | /clone="QmoA-13657" | | |
| | /sex="male" | | |
| | /tissue_type="medulla oblongata" | | |
| | /clone_lib="macaque brain cDNA library QmoA" | | |
| | /dev_stage="adult" | | |
| | 6..653 | | |
| | /codon_start=1 | | |
| | /product="hypothetical protein" | | |
| | /protein_id="BAC41746.1" | | |
| | /db_xref="gi:26449237" | | |
| | /translation="MPARNRPPLVSLVLITMASVCPCYCVESSETEPAOVGNPMKLK CISMKKEVEATIVWFVRPEGSKDLITRYNHGHDEVSPQGRLOMGKSLOLD VSILVKLVNTLNDSGLYTCNVREGEFPEAHKRPVFTTRILPIRVTEEAGEDSTVSASB | | |
| CDS | | | |

| ORIGIN | Query Match | 71.8%; | Score 462.8; | DB 9; | Length 3296; |
|------------|---|---|-------------------|-----------|-----------------|
| | Best Local Similarity | 58.4%; | Pred. No. 1e-105; | | |
| | Matches 377; | Conservative 145; | Mismatches 123; | Indels 0; | Gaps 0; |
| Qy | 1 | ATGCCNCNNTTAAAYMGNYNTTTCNCYTCNGCNSMNTNGTNTYNTATNTAATYAGGTGTT | 60 | | |
| Db | 6 | ATGCCCTCCTTCAACAAGATTGTTTCCCTCGTGTCTCTCGTGTCTATCTACTGCGCCAGT | 65 | | |
| Qy | 61 | GTNTGYTTTCNCGTNGTGYTNGARGTNCNWSNGARACNGRGCNTNCARGNAAYCCN | 120 | | |
| Db | 66 | GTCTGCTTCCCGGTGTGTGTGAAGTCCCTCCGAGACGAGCGCTGTGTCAAGGCAACCC | 125 | | |
| Qy | 121 | ATGAARATYMNNGTGYATNTMSGTGATGAABNGNARARGTNGARGCNACNAGTNGTN | 180 | | |
| Db | 126 | ATGAAGCTGCGCGTCACTCTCTGATGAAGAAGAGAGGTGGAGGCCACACAGTGTG | 185 | | |
| Qy | 181 | GARTGTTTAAVMNCCNGARGGNGNAAARGATYTTTNTATNTAYGARTAYMNAAYGN | 240 | | |
| Db | 186 | GAATGTCTTCAACAGCCCGAGGGGGTAAAGATTCTTATCTACGATATCGGAATGCG | 245 | | |
| Qy | 241 | CATCARBARGTNGARNSNCCCTTTCARGMNGNTNCTARTGGAAYGGMWSMAAGATYTN | 300 | | |
| Db | 246 | CACCAAGAGGTAGAGAACCCCTTTCAGGCGGCTGTGAGATGTGAGCAAGACGACTG | 305 | | |
| Qy | 301 | CARGAYGTNMSNATNACGTNTYNTAAAYGTNACNTYNTAAYGAYMNGNYTNTAAYCNTGY | 360 | | |
| Db | 306 | CAGGACGTGTCTCATCTGTGCTCAAGTCACTGTGAACGACTCTGGCGCTTCACTGCG | 365 | | |
| Qy | 361 | AAATGTNMSNNGARTTYGARTTYGARGCNCAYMNCNTTGYNTAARACNACMNGNTN | 420 | | |
| Db | 366 | AAATGTGCCCGGAGTGTGAGTGTGAGCGCATGCGCCCTTGTGAAAGACACACCGCGTG | 425 | | |
| Qy | 421 | ATNCCNTTNGNGTNCNGARGCNGGNGNGARGATYTTTACMSNNGTNGTWSNGARATN | 480 | | |
| Db | 426 | ATCCCACTAAGAGTCAACCGAGGAGGCGGAGAGACCTTCACTCTGTGTGTCTCGAANAATC | 485 | | |
| Qy | 481 | ATGATGATYATNTYNTNGTNTTYYTNAACNTYNTGTNTYNTATNGARATGATNTAYTGY | 540 | | |
| Db | 486 | ATGATGTACATCTTCTGTGTCTTCTCTCACTTGTGGCTGCTCATGTGATGATATATTGC | 545 | | |
| Qy | 541 | TATWGNAAARGTMSNARGCNGARGCNGCNCARBARAAYGCMNSGATAYTNTGNCN | 600 | | |
| Db | 546 | TACGAAAGGCTCAAAAGCCGAAAGACGAGCCCAAGAAACGCGTCTGACTCACTTGTGCC | 605 | | |
| Qy | 601 | ATNCCNMSNGARAAYARGAAAYWSNGCNGTNCNGTNGARGAR | 645 | | |
| Db | 606 | ATCCCATCTGAGAACAGAGAACTCTGCACTACAGTACGAGGAA | 650 | | |
| RESULT 13 | | | | | |
| LOCUS | BC070899 | 3910 bp | mRNA | linear | ROD 28-SEP-2004 |
| DEFINITION | Rattus norvegicus sodium channel, voltage-gated, type IIf, beta, | | | | |
| ACCESSION | BC070899 | | | | |
| VERSION | BC070899.1 | GI:47477789 | | | |
| KEYWORDS | MGC. | | | | |
| SOURCE | Rattus norvegicus (Norway rat) | | | | |
| ORGANISM | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | |
| | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; | | | | |
| | Rattus | | | | |
| REFERENCE | 1 (bases 1 to 3910) | | | | |
| AUTHORS | Strauhuber, R.L., Feingold, B.A., Grouse, L.H., Derge, J.G., | | | | |
| | Klauser, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., | | | | |
| | Altschul, S.F., Zeeberg, B., Buettow, K.H., Schaefer, C.F., Bhat, N.K., | | | | |
| | Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heien, F., | | | | |
| | Diatchenko, L., Marisina, K., Farmer, A.A., Rubin, G.M., Hong, L., | | | | |
| | Scapleorn, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., | | | | |
| | Schaefer, T.E., Brownstein, M.J., Udell, T.B., Toshiyuki, S., | | | | |


```

/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="MGC:56857 IMAGE:6308278"
/cisue_type="Olfactory epithelium, neonatal mouse,
C57BL/6"
/clone_lib="NIH_MGC_129"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
1..4169
/gene="Scn3b"
/notes="synonym: Scn3"
/db_xref="locusid:235281"
/db_xref="MGI:191882"
308..955
/gene="Scn3b"
/codon_start=1
/product="Scn3b protein"
/protein_id="AAH53919.1"
/db_xref="GI:32172798"
/db_xref="MGI:191882"
/translaction="MPAFNRLPLASIVLIYVAVCPVCEVPESETEBAVQSNMKLR
CISCMKBEVATTVVYFVRPEGKDFLIYEVNHOEVSPEGRLOWNSKLDL
VSIIVLVITLNDGLTQVNSREPEFAHRRFVKTTRLLIPIRYTEAGDEPTSVSEI
MYTILVFLTILMLETIMTYCYRKVSKAEBAQENASDYLALPSNKENSVPVEE"
443..673
/gene="Scn3b"
/notes="IGV; Region: Immunoglobulin V-Type"
/db_xref="CDD:smart00406"

misc_feature
43..673

ORIGIN
Query Match 71.5%; Score 461.2; DB 10; Length 4169;
Best Local Similarity 58.3%; Pred. No. 2.8e-105;
Matches 376; Conservative 145; Mismatches 124; Indels 0; Gaps 0;

QY 1 ATGCNCGNTTAAVMGVNTTTCNTYNGCNMSNTYNTNATNTATYTGGTGKMSN 60
DB 308 ATGCTGCTTCAACAGATGCTTCCCTACCTCTCTAGTCTCATCTACTGCTCAGA 367
QY 61 GTNTGTYTTCNGTNTGTYGTNGARGTNCNMSNGARACNGARGCNGTNCARGNAAYCN 120
DB 368 GTCTGCTTCCCTGTGTGTAGAGTACCTCGAGACAGAGCCGTGCAGGCAATTCC 427
QY 121 ATGAARTYTNMGNTGYATNWSVTGYATGAARMNGARGGTNGARBCNACNACNGTNGTN 180
DB 428 ATGAAGCTGAGATCATCTCTGCATGAAGAGGAGGAGGTGAGGCCACCATGTAGTG 487
QY 181 GARTGTTTATYMGNCNGARGGNGNARAGATTTTNTATNTATYGAATATMGNAAYGN 240
DB 488 GAGTGTTCACAGGCTGAGGCGGTAAAGATTTCTATATATGATATCGAATATGC 547
QY 241 CAYCARGAGTNGARWNSCNTTYCARGNMGNTNCACTGGAAYGWSNARAGAYTN 300
DB 548 CACCAAGAGGTGAGAGAGCCCTTCCAAAGTGTCTGCAGTGAATGGAGCAAGACCTG 607
QY 301 CARGAYGTMSNATNACNGTNTYNAAYGTNAANTYNAAYGAYWSNGANTYNTAYACNTGY 360
DB 608 CAGAGCATATCATCATCTGTCTCATATGTCACTGATATGACTGTGGCCTTACACATGT 667
QY 361 AAVGTMSNMGNGARTTYGATTYGARGCNCAYMNCNTTYGTNAARACNACMGNTYN 420
DB 668 AATGTTCAGGGAGTGTGAGTGAAGCACACCGCCCTTGTGAGACCAACAAGACTA 727
QY 421 ATNCGNTYTNMGNTNACGARGCNGARGCNGARGAYTTTACNWSNGTNGTWSNGARATN 480
DB 728 ATACCCCTCGAGACACTGAGAGAGCGGAGAGAACTTCACTCGGTGTCTCGGAATC 787
QY 481 ATGATGTATATNTYNTNTNTTNTTNTNACNTTNTGTYNTNTATNAGARATNTAYTGY 540
DB 788 ATGATGTATACCTCTCTGCTCTCTCATCTGTGTGCTGTTATYTAGATATCTATTTGC 847
QY 541 TAYMGNAAGTMSNARGCNGARGCNGCNCARBARAYGCMWSGATYATYTNCGN 600

```

```

DB 848 TACGAAAGTCTCTTAAGCCGAGAGAGGACGCTCAGAAATGCTCTGACTTCTCT 907
QY 601 ATNCCNWSNGARAAAYARAGARAAAYWSNGCNGTNCNGTNGARGAR 645
DB 908 ATCCCTTCAGAGAACAGAGAACTCTGTGTGTACCCGTGAGAGAA 952

```

Search completed: April 7, 2005, 22:37:09
Job time : 3012 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2005, 19:47:17 ; Search time 434 Seconds

(without alignments)
8797.772 Million cell updates/sec

Title: US-09-977-579-2-REV

Sequence: 1 ATGCCGNCNTTYAAYMGNTY.....SNGCNGTNCNGTNGANGAR 645

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1990s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------------|--------------------|
| 1 | 466 | 72.2 | 1261 | 3 AAC67837 | Aac67837 Human bet |
| 2 | 466 | 72.2 | 1261 | 10 ADB78651 | Adb78651 Human ion |
| 3 | 466 | 72.2 | 1261 | 10 ACFS7870 | Acfs7870 Human ion |
| 4 | 466 | 72.2 | 1510 | 4 AAF84146 | Aaf84146 Human nov |
| 5 | 466 | 72.2 | 4052 | 6 ABA93727 | Aba93727 Human sig |
| 6 | 462.8 | 71.8 | 645 | 5 AAC90601 | Aac90601 Rat sodiu |
| 7 | 462.8 | 71.8 | 2220 | 3 AAC67836 | Aac67836 Rat betaf |
| 8 | 462.8 | 71.8 | 2632 | 5 AAC90602 | Aac90602 Rat sodiu |
| 9 | 462.8 | 71.8 | 3108 | 5 AAC90600 | Aac90600 Rat sodiu |
| 10 | 424.8 | 65.9 | 1045 | 4 AAK52345 | Aak52345 Human pol |
| 11 | 412.4 | 63.9 | 978 | 4 AAH98320 | Aah98320 Human EST |
| 12 | 412.4 | 59.9 | 978 | 13 ADS11487 | Ads11487 Human the |
| 13 | 384 | 63.9 | 1195 | 5 AAS86764 | Aas86764 DNA encod |
| 14 | 363.2 | 56.7 | 953 | 13 ADS10151 | Ads10151 Human the |
| 15 | 301 | 46.7 | 471 | 2 AAV86895 | Aav86895 EST clone |
| 16 | 285.2 | 44.2 | 3531 | 5 AAS86763 | Aas86763 DNA encod |
| 17 | 163 | 25.3 | 621 | 5 AAS86762 | Aas86762 DNA encod |
| 18 | 158.4 | 24.6 | 1335 | 10 ACFS7868 | Acfs7868 Human SCN |
| 19 | 158.4 | 24.6 | 1335 | 12 ADQ19428 | Adq19428 Human sof |
| 20 | 158 | 24.5 | 657 | 5 AAC90603 | Aac90603 Rabbit so |

| | | | | | |
|----|-------|------|------|-------------|---------------------|
| 21 | 157.2 | 24.4 | 1414 | 10 ADB78641 | Adb78641 Human ion |
| 22 | 156.8 | 24.3 | 1414 | 8 ABZ23837 | Abz23837 Human vol |
| 23 | 156.8 | 24.3 | 1414 | 10 ADB78640 | Adb78640 Human ion |
| 24 | 152.2 | 23.6 | 1490 | 6 ABK63747 | Abk63747 Rat seque |
| 25 | 152.2 | 23.6 | 1490 | 10 ADB52855 | Adb52855 Primary r |
| 26 | 152.2 | 23.6 | 1490 | 12 AD009498 | Ado09498 Rat sodiu |
| 27 | 105 | 16.3 | 407 | 4 ABA08942 | Abao8942 Human vol |
| 28 | 105 | 16.3 | 407 | 4 AAK53329 | Aak53329 Human pol |
| 29 | 89.6 | 13.9 | 358 | 10 ACD98234 | Acdd98234 Human col |
| 30 | 86 | 13.3 | 855 | 10 ADG15041 | Adg15041 Human SEC |
| 31 | 81.2 | 12.6 | 974 | 6 AAD29622 | Aad29622 Human bet |
| 32 | 73.8 | 11.4 | 850 | 4 AAF5675 | Aaf5675 Rat sodiu |
| 33 | 73.8 | 11.4 | 850 | 4 AAF30659 | Aaf30659 Sodium ch |
| 34 | 72.6 | 11.3 | 2000 | 8 ADA71938 | Ada71938 Rice gene |
| 35 | 69.8 | 10.8 | 2000 | 8 ADA71938 | Ada71938 Rice gene |
| 36 | 57.8 | 9.0 | 162 | 12 ADM66944 | Adm66944 Human adi |
| 37 | 57.8 | 9.0 | 3583 | 12 ADQ23651 | Adq23651 Human sof |
| 38 | 56.6 | 8.8 | 842 | 12 ADM66943 | Adm66943 Murine ad |
| 39 | 54.8 | 8.5 | 4590 | 5 AAH24065 | Aah24065 Yeast AOD |
| 40 | 52.6 | 8.2 | 6217 | 6 ABL32938 | Ab132938 Human imm |
| 41 | 51 | 7.9 | 5454 | 3 AAA70236 | Aaa70236 Plasmodiu |
| 42 | 49.6 | 7.7 | 2415 | 4 AAC84367 | Aac84367 Human Zac |
| 43 | 49.6 | 7.7 | 2415 | 10 ABX93334 | Abx93334 Degenerat |
| 44 | 49.4 | 7.7 | 633 | 10 ADB68844 | Adb68844 Minorily/ |
| 45 | 48.6 | 7.5 | 5059 | 2 AAX84332 | Aax84332 Stealth v |

ALIGNMENTS

| | |
|----------|---|
| RESULT 1 | |
| AAC67837 | standard; cDNA; 1261 BP. |
| ID | AAC67837 |
| AC | AAC67837; |
| XX | |
| DT | 15-FEB-2001 (first entry) |
| XX | |
| DE | Human beta3 cDNA. |
| XX | |
| KW | Human; beta sub-unit; beta3; analgesic; anticonvulsant; |
| KW | cerebroprotective; vasotropic; cardiant; nootropic; cytostatic; |
| KW | dermatological; gene therapy; voltage-gated sodium channel; pain; |
| KW | epilepsy; stroke; ischaemia; heart disease; Jacobsen Syndrome; |
| KW | familial nonchromaffin paraganglioma; phenylketonuria; |
| KW | Charcot Marie Tooth disease; ss. |
| XX | |
| OS | Homo sapiens. |
| XX | |
| PN | W0200063367-A1. |
| XX | |
| PD | 26-OCT-2000. |
| XX | |
| PF | 24-FEB-2000; 2000WO-BP001783. |
| XX | |
| PR | 15-APR-1999; 99US-0129473P. |
| XX | |
| PA | (WARN) WARNER LAMBERT CO. |
| XX | (UYCA-) UNIV CAMBRIDGE TECH SERVICES LTD. |
| PI | Cox P, Dixon A, Jackson A, Morgan K; |
| XX | |
| DR | WPI; 2000-665241/64. |
| XX | P-PSDB; AAB36002. |
| PT | Novel nucleic acids encoding a beta-3 subunit from a voltage-gated sodium |
| PT | channel, and their corresponding polypeptides, useful for detecting and |
| PT | treating sodium channel-associated conditions, e.g. pain, epilepsy and |
| PT | stroke. |
| PS | Claim 10; Page 70-71; 88pp; English. |
| XX | |
| XX | The present sequence is given in the claims of a specification relating |

CC to a novel family of beta sub-unit proteins from a voltage-gated sodium
CC channel. Human and rat beta sub-units, which have been collectively
CC identified as *brev2*, have been isolated. The polynucleotides and
CC polypeptides are useful for screening for agonists and antagonists of
CC sodium channels. The agonists, antagonists, proteins and nucleic acids
CC may be used diagnosing of treating diseases or conditions associated with
CC voltage-gated sodium channels, e.g. pain, epilepsy, stroke, ischemia,
CC heart disease, Jacobson Syndrome, Familial Nonchronic Pain Paraganglioma,
CC Phenylketonuria and Charcot Marie Tooth disease

SQ Sequence 1261 BP; 263 A; 366 C; 372 G; 260 T; 0 U; 0 Other;

| | | | | |
|-----------------------|------------------|--------------------|----------|-------------|
| Query Match | 72.2% | Score 466 | DB 3 | Length 1261 |
| Best Local Similarity | 58.8% | Pred. No. 1.6e-117 | | |
| Matches 379 | Conservative 121 | Mismatches 121 | Indels 0 | Gaps 0 |

[illegible]

| | | | |
|----|-----|---|------|
| QY | 241 | CAYCARGARGNTNAGRMSNCNTTTCARAGGMMNTNCAFTGGAAVCGMSNNAARGAAYTN | 300 |
| Db | 616 | CACCAAGAGTGGAAGGCCCTTTTCAGGGGCGCTCGAGTGGATATGACAGCAAGACCTGG | 675 |
| QY | 301 | CARGAYGTNMSNATNACNGTNTYTNAAYGTNACNYTNAAGVAYMSNGNYTNTAYACNTGY | 360 |
| Db | 676 | CAGACAGTGTCCATCATCTGTGCTCAACGTCACTCTGAACGACTGTGGCCTCTACACTGG | 735 |
| QY | 361 | AAVGTNMSMNGANTTYGARTTYGARVGCNCAVMGNCNTTYGTNAAAPACNACMNGAYTN | 420 |
| Db | 736 | AAATGTGTCGCCGGAAGTTGAGTTTGAGGCGCAATCGCCCTTTGTGAAGACGACGCCGCTG | 795 |
| QY | 421 | ATNCCNNTNMGGTNAACNGARBARCGCNGSNGARGATTYTAACMSNCTNNGSNNGARATN | 480 |
| Db | 796 | ATCCCCCTPAAGGTACCCGAGAGGCTGAGAGGAACTTCACTCTGTGTCTTAGAAATC | 855 |
| QY | 481 | ATGATGTATVATYNTVNTGTNTTYYTNACNNTNTGGTYTNTATNAPATGATNTATYTG | 540 |
| Db | 856 | ATGATGTATCATCTCTGTGTCCTTCCACACCTGTGCTGCTCATGAGATGATATATTTGC | 915 |
| QY | 541 | TAYMGNAARGTNSNAARGCNGARGARGCNGCNCARGAPAAVGCNMSNGAYTAYTYTNGCN | 600 |
| Db | 916 | TACAGAAAGGTCTCAAAAAGCCGAAGAGGACAGCCCAAGAAAACGCGTCACTACCTTGGCC | 975 |
| QY | 601 | ATNCCNMSNGARAAAYARBARAAYMSNGCNTNCCNGTNGARGAR | 645 |
| Db | 976 | ATCCCATCTTGAGAACAGGAACCTGTGGGTATCCAGTGAGAGAA | 1020 |

RESULT 2
ADB78651
ID ADB78651 standard; cDNA; 1261 BP.

| | |
|----|---------------------------|
| AC | ADB/8651; |
| XX | |
| DT | 04-DEC-2003 (first entry) |

DE Human ion channel subunit CDNA mutant SCN1A^r exon 1 SEQ ID NO:22.

KW ss; gene; mutant; ion channel; ion channel subunit; ICS; nootropic;
 neuroprotective; inotropic; antiarrhythmic; antidiarrhoeic; antimigraine;
 KW

KW antidepressant; antiparkinsonian; neuroleptic; tranquilizer; analgesic;
KW nephrrotropic; antidiabetic; ophthalmological; epilepsy;
KW ion channel dysfunction; human.

KW ion channel dysfunction; human.

OS Synthetic.

Homo sapiens.

PN WO2003008574-A1

PD 30-JAN-2003

PF 08-JUL-2002; 2002WO-AU000910.

PR 18-JUL-2001; 2001AU-00006452.

PR 13-MAY-2002; 2002AU-000022292.

PA' (BION-) BIONOMICS LT

XXXX / XXXXXXX N. W.
XXXX / XXXXXXX N. W.
XXXX / XXXXXXX N. W.
XXXX / XXXXXXX N. W.

PI Berkovic SF. Scheffer IE; Mulley JC, Harkin LM, D

XX
NR
WPT: 2003-239332/23.

Identifying predilection to an ion channel dysfunction. Such as

periodic paralysis, cardiac arrhythmias, migraine, Alzheimer's disease, schizophrenia, anxiety and depression by detecting encoding-gene

PT mutation events.

PS Claim 6; SEQ ID NO 22; 106pp; English.

CC The invention relates to a novel method for identifying a subject
CC predisposed to a disorder associated with ion channel dysfunction. The
CC method comprises ascertaining if at least one of the genes encoding ion
CC channel subunits (ICs) has undergone a mutation event so that a cDNA
CC derived from the subject has any of 134 nucleotide sequences. The method
CC of the invention has neurotropic, neuroprotective, inotropic, antihypertetic,
CC antiarrhythmic, antimigraine, antidepressant, antiparkinsonian,
CC neuroleptic, tranquiliser, analgesic, nephroprotective, antidiabetic, and
CC ophthalmological activity. A polynucleotide of the invention acts as an
CC ion channel agonist, or ion channel antagonist. The methods, isolated
CC nucleic acids, polypeptides, antibody, selective agonist, antagonist or
CC modulator of an ion channel, cells and genetically modified non-human
CC animal, are useful for the diagnosis and treatment of epilepsy and/or a
CC disorder associated with ion channel dysfunction, such as hyper- or hypo-
CC kalemic periodic paralysis, myotonia, malignant hyperthermia,
CC myasthenia, cardiac arrhythmias, episodic ataxia, migraine, Alzheimer's
CC disease, Parkinson's disease, schizophrenia, hyperekplexia, anxiety,
CC depression, phobic obsessive symptoms, neuropathic pain, inflammatory
CC pain, chronic/acute pain, Bartter's syndrome, polycystic kidney disease,
CC Dent's disease, hypernephritic hypoglycaemia of infancy, cystic
CC fibrosis, congenital stationary night blindness and total colour
CC blindness. The present sequence represents a mutant cDNA of the
CC invention. The sequence data for this patent is not represented in the
CC printed specification, but was obtained in electronic format directly
CC from http://www.wipo.int/publ/published_pat_sequences.

SQ Sequence 1261 BP; 263 A; 365 C; 372 G; 261 T; 0 U; 0 Other;

| | | | | |
|-----------------------|------------------|--------------------|----------|-------------|
| Query Match | 72.2% | Score 466 | DB 10 | Length 1261 |
| Best Local Similarity | 58.8% | Pred. No. 1.6e-117 | | |
| Matches 379 | Conservative 145 | Mismatches 121 | Indels 0 | Gaps 0 |

QY 1 ATGCCNGCNTTYAAYMGNTNTTYCCNYTNGCNSNYTNGTNTNATNTAYTGGGTWSN 60

Db 376 ATGCCCTGCCTTCAATAGATTGTTCCCCCTGGCTTCTCTCGTGTATATCTACTGGTCAGT 435

61 GTNTGYTTYCCNGTNTGYGTNGARGTNCNWSNGARACNGARGCNGTNCARGGNAAYCCN 120

Db 436 GTCTGCTTCCCTGTGTGTGGAAGTCCCTCGGAGACGGAGGCCCTGCAGGGCAACCC 495

121 ATGAARYTNMGNTGYATNWSNTGYATGARMGNGARGRGNGARGCNCACNGTNGTN 180

[illegible]

| XX | RESULT 3 |
|----|--|
| XX | ACF57870 |
| XX | ID ACF57870 standard; cDNA; 1261 BP. |
| XX | ACF57870; |
| XX | 15-JAN-2004 (first entry) |
| XX | Human SCN3B protein encoding cDNA. |
| XX | SCN1A; sodium channel type 1 alpha-subunit; anticonvulsant; analgesic; |
| XX | neuriprotective; anesthetic; cytostatic; cerebroprotective; cardiant; |
| XX | hypotensive; gene therapy; SCN3B; human; gene; ss. |
| XX | Homo sapiens. |
| XX | OS |
| XX | Key |
| XX | Location/Qualifiers |
| XX | 376..1023 |
| XX | CDS |
| XX | FT |
| XX | FT |
| XX | /tag= a |
| XX | /product= "SCN3B" |
| XX | MO2003072751-A2. |
| XX | PD |
| XX | 04-SEP-2003. |
| XX | XX |
| XX | 25-FEB-2003; 2003MO-US006010. |
| XX | XX |
| XX | 25-FEB-2002; 2002US-0359382P. |
| XX | XX |
| XX | PA |
| XX | (UTVA-) UNIV VANDERBILT. |
| XX | George AL, Loezin C; |
| XX | PI |
| XX | WPI; 2003-712725/67. |
| XX | DR |
| XX | P-PsDB; ABR63183. |
| XX | Recombinantly expressed sodium channel type 1 alpha subunit, useful in |

| | | |
|----|--|---|
| Pt | | screening for modulators, for treating e.g. epilepsy. |
| Xx | | |
| P8 | | |
| Xx | Disclosure; | Page 145-147; 176pp; English. |
| Cc | The invention relates to a recombinantly expressed and isolated human | |
| Cc | SCNIA (sodium channel type I alpha-subunit) (I). (II), optionally | |
| Cc | incorporated into a cell, is used to screen for specific modulators, | |
| Cc | potentially useful as anticonvulsant, anti epileptic, neuroprotective, | |
| Cc | anesthetic and/or anaesthetic agents, e.g. for treating severe myoclonic | |
| Cc | epilepsy of infancy, stroke, cardiac arrest, hyperkalemic paralysis, and | |
| Cc | motor endplate diseases, hypertension, congestive heart failure and | |
| Cc | mucular dystrophy also to treat cancer (SCNIA is expressed in prostatic | |
| Cc | and metastatic cancer cell lines). These activities can also be provided | |
| Cc | by gene therapy vectors that express (I) or the modulators. The | |
| Cc | modulators, also antibodies directed against (I), are used to detect | |
| Cc | sodium channel polypeptides. The present sequence represents a human | |
| Cc | SCN3B protein encoding cDNA | |
| Xx | | |
| SQ | Sequence | 1261 BP; 263 A; 366 C; 372 G; 260 T; 0 U; 0 Other; |
| | Query Match | 72.2%; Score 466; DB 10; Length 1261; |
| | Best Local Similarity | 58.8%; Pred. NO. 1.6e+117; |
| | Matches 379; Conservative 145; Mismatches 121; Indels 0; Gaps 0; | |

[illegible]

```

AC  AAF84146;
XX
XX 07-SEP-2001 (first entry)
XX
XX Human novel sodium channel beta1-like subunit encoding cDNA.
DE
XX Sodium channel; sensory neurone specific channel, beta1-like subunit;
XX SNS; therapeutic; pain; analgesic; ss.
XX
XX Homo sapiens.
OS
XX Key Location/Qualifiers
XX CDS 213..860
XX FT /*tag= a
XX FT /product= "sodium channel beta1-like subunit"
XX
XX MO200144293-A2.
XX
XX 21-JUN-2001.
XX
XX 14-DEC-2000; 2000MO-GB004802.
XX
XX 17-DEC-1999; 99GB-00029970.
XX
XX (GLAX ) GLAXO GROUP LTD.
XX
XX Plumptre M, Powell AJ, Sansau P;
XX
XX WPI; 2001-398129/42.
XX
XX P-PSDB; AAB85206.
XX
XX Novel sub-unit for voltage-gated sodium channel proteins for producing
XX PT agents useful for treating pain.
XX
XX Claim 4; Page 29-30; 31pp; English.
XX
XX The invention provides a novel beta1-like sub-unit for voltage-gated
XX sodium ion channel polypeptide, specifically a sensory neurone specific
XX channel (SNS) subunit. The novel beta1-like subunit is useful for
XX producing a therapeutic agent which is useful treating pain in a patient.
XX The subunit can be expressed by standard recombinant methodology. The
XX present sequence represents a human novel sodium channel beta1-like
XX subunit encoding cDNA
XX
XX Sequence 1510 BP; 350 A; 410 C; 414 G; 336 T; 0 U; 0 Other;
SQ
Query Match 72.2%; Score 466; DB 4; Length 1510;
Best Local Similarity 58.8%; Pred. No. 1.8e-117;
Matches 379; Conservative 145; Mismatches 121; Indels 0; Gaps 0;
QY 1 ATGCCGACNTTAAAGNNTTCCNTGTCGCTTCTCTGCTTATCTAGGTCGATG 60
DB 213 ATGCCGACNTTAAAGNNTTCCNTGTCGCTTCTCTGCTTATCTAGGTCGATG 272
QY 61 GTTGTGTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
DB 273 GTTGTGTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 332
QY 121 ATCAATTTTGGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 180
DB 333 ATCAATTTTGGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 392
QY 181 GATTGTGTTAAAGNNTTCCGCTGCTTCTCTGCTTCTCTGCTTCTCTGCTTCTCT 240
DB 393 GATTGTGTTAAAGNNTTCCGCTGCTTCTCTGCTTCTCTGCTTCTCTGCTTCTCT 452
QY 241 CAYCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
DB 453 CAYCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 512
QY 301 CAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
DB 513 CAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 572

```

```

QY 361 AAATGTTTAAAGNNTTCCGCTGCTTCTCTGCTTCTCTGCTTCTCTGCTTCTCT 420
DB 573 AAATGTTTAAAGNNTTCCGCTGCTTCTCTGCTTCTCTGCTTCTCTGCTTCTCT 632
QY 421 ATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
DB 633 ATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 692
QY 481 ATGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
DB 693 ATGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 752
QY 541 TATGTTTAAAGNNTTCCGCTGCTTCTCTGCTTCTCTGCTTCTCTGCTTCTCT 600
DB 753 TATGTTTAAAGNNTTCCGCTGCTTCTCTGCTTCTCTGCTTCTCTGCTTCTCT 812
QY 601 ATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 645
DB 813 ATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 857

RESULT 5
ABA93727
ID ABA93727 standard; cDNA; 4052 BP.
XX
XX ABA93727;
XX
XX 30-APR-2002 (first entry)
XX
XX Human signal transduction cDNA clone amy2_2f18.
XX
XX Human; foetal brain; foetal kidney; melanoma; testis; amygdala;
XX gene therapy; ss.
XX
XX Homo sapiens.
XX
XX MO200198454-A2.
XX
XX 27-DEC-2001.
XX
XX 25-APR-2001; 2001MO-IB002050.
XX
XX 25-APR-2000; 2000US-0199380P.
XX
XX (GERM-) GERMAN HUMAN GENOME PROJECT.
XX
XX Wiemann S;
XX
XX WPI; 2002-055860/07.
XX
XX P-PSDB; ABB05689.
XX
XX Human cDNA sequences and clones derived from human fetal brain, fetal
XX kidney, melanoma, testis and amygdala cDNA libraries, useful in genetic
XX screening and therapy.
XX
XX Claim 1; Page 174-175; 611pp; English.
XX
XX The present invention describes assemblages and computer readable media
XX comprising novel human cDNA sequences and clones derived from human
XX foetal brain, foetal kidney, melanoma, testis and amygdala cDNA
XX libraries. ABA93702 to ABA93766 represent human cDNA sequences from the
XX present invention which encode the proteins given in ABB05662 to
XX ABB05729. The human cDNA sequences and clones can be used in gene
XX therapy. The clones may be used in a variety of applications, for example
XX they may be used in profiling assays, for providing large arrays of human
XX genetic material for implementing large-scale screening strategies and
XX for treating diseases via gene therapy procedures
XX
XX Sequence 4052 BP; 994 A; 1033 C; 1036 G; 989 T; 0 U; 0 Other;
SQ
Query Match 72.2%; Score 466; DB 6; Length 4052;
Best Local Similarity 58.8%; Pred. No. 3.8e-117;

```

Matches 379; Conservative 145; Mismatches 121; Indels 0; Gaps 0;

QY 1 ATGCCNCGNTTYAAYMGNTYNTTYCCNYTNGCMNSNTNGTNTATNTATYTGCTGMSN 60
DB 804 ATGCTGCTCTCAATAGATGTTTCCCTGCTCTCTGCTGCTATCTACTGCTGCTAGT 863
QY 61 GTTNGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
DB 864 GTTNGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 923
QY 121 ATGAATYTMNGTNTGATYTMNSGTATGTAARMNGARGARGNGARGNACNACNGTNGT 180
DB 924 ATGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 983
QY 181 GARTGTTTATYTMNGCNGARGNGARGNARGAATYTTTATNTATYGAATYTMNSAAYGN 240
DB 984 GAATGTTCTACAGGCGCGAGGCGGTAAAGATTCTTATTTACGAGTATCGGAATGCT 1043
QY 241 CAYCARGARGTNGARGMSNCTTTCARGMNGTNTGARGAAYGMSNAARGAAYTN 300
DB 1044 CACGAGAGGTGAGAGCCCTTTCAGGGGCGCTGCGATGGAATGCGAGAGGAGCTG 1103
QY 301 CARGAYTMNSATNACNGTNTATYTMNSATNTATNTATNTATNTATNTATNTATNTAT 360
DB 1104 CAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1163
QY 361 AAYGTMSNMGNGARTTYGARTTYGARGCNCAYMGNCNTTYGTNAARACNACMNGYTN 420
DB 1164 AATGTGCCCGGAGTTTGAATTTGAGGCGCATCGGCCCTTGTGTAAGACAGCGGCTG 1223
QY 421 ATNCCNTNMNGNACNGARGCNGARGARGAATYTTACMNSNGTNGTMSNARGATN 480
DB 1224 ATCCCCCTTAAGATCACCGAGAGCGTGAAGAGACTTCACTGCTGCTGCTGCTGCTG 1283
QY 481 ATGATGTAATNTYNTNGTNTTYTNAACNTNTGTYNTNTATNGARATGATNTATYTG 540
DB 1284 ATGATGTAATCTCTTCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1343
QY 541 TAYWGNARGTMSNARGCNGARGCNGCNGCARGAAYGCMNSNGATYATYTMNGN 600
DB 1344 TACGAGAAAGTCTCAAAAGCGGAGAGCGACCCCAAGAAAAGCGCTGCTGACTACCTG 1403
QY 601 ATNCCMNSNGARAAYARGARAAYMSNGCNGTNGTNGARGAR 645
DB 1404 ATCCATCTGAGAACAGAGAACTCTGCGGTACCACTGAGAGAA 1448

RESULT 6
AAC90601
ID AAC90601 standard; cDNA, 645 BP.
XX AAC90601;
AC AAC90601;
DT 13-MAR-2001 (first entry)
XX
DE Rat sodium channel beta3 protein Alrx94h5 partial coding sequence.
XX Rat; sodium channel beta3 protein; Alrx94h5; pain; sleep disorder;
KM neurodegenerative disorder; mood disorder; muscle contraction; ss.
XX
OS Rattus sp.
XX
PN W0200069912-Al.
XX
PD 23-NOV-2000.
XX
PE 12-MAY-2000; 2000WO-US013144.
XX
PR 14-MAY-1999; 99US-0134198P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Curtiss RAJ;

XX
DR WPI: 2001-122743/13.
XX
PT New rat sodium channel beta-3 subunit gene isolated from a rat dorsal
PT root ganglion cDNA library for use in chromosome mapping, forensic
PT medicine, monitoring clinical trials and therapeutics.
XX
XX
XX Claim 1; Page 40-41; 145pp; English.
XX
CC The present invention provides the protein and coding sequences of the
CC rat sodium channel beta3 protein, designated Alrx94h5. This protein is
CC involved in the generation of pain and other sensory or perceptible nerve
CC impulses, in the establishment and endurance of mood, neurodegenerative
CC and sleep disorders, and in the control of muscle contraction, including
CC movements such as the heartbeat, digestion and vascular tone. The
CC sequences can be used in predictive medicine, screening and diagnostic
CC assays, and in pharmacogenomics
XX
SQ Sequence 645 BP; 155 A; 154 C; 181 G; 155 T; 0 U; 0 Other;
Query Match 71.8%; Score 462.8; DB 5; Length 645;
Best Local Similarity 58.4%; Pred. No. 7.4e-117;
Matches 377; Conservative 145; Mismatches 123; Indels 0; Gaps 0;

QY 1 ATGCCNCGNTTYAAYMGNTYNTTYCCNYTNGCMNSNTNGTNTATNTATYTGCTGMSN 60
DB 1 ATGCTGCTCTCAATAGATGTTTCCCTGCTCTCTGCTGCTATCTACTGCTGCTAGT 60
QY 61 GTTNGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
DB 61 GTTNGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
QY 121 ATGAATYTMNGTNTGATYTMNSGTATGTAARMNGARGARGTNGARGCNCNACNGTNGT 180
DB 121 ATGAAGCTGAGTGCATCTCTGCTGATGAAGAGGAGGAGGTGAGGCGCACCATGCTG 180
QY 181 GARTGTTTATYTMNGCNGARGCNGARGAATYTTTATNTATYGAATYTMNSAAYGN 240
DB 181 GARTGTTTATYTMNGCNGARGCNGARGAATYTTTATNTATYGAATYTMNSAAYGN 240
QY 241 CAYCARGARGTNGARGMSNCTTTCARGMNGTNTGARGAAYGMSNAARGAAYTN 300
DB 241 CACGAGAAAGTGAAGACCCCTTTCAGGGCGGTGAAGATTTCTTATATYAGTATCGGAATG 240
QY 301 CARGAYTMNSATNACNGTNTATYTMNSATNTATNTATNTATNTATNTATNTATNTAT 360
DB 301 CAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
QY 361 AAYGTMSNMGNGARTTYGARTTYGARGCNCAYMGNCNTTYGTNAARACNACMNGYTN 420
DB 361 AATGTGCCCGGAGTTTGAATTTGAGGCGCATCGGCCCTTGTGTAAGACAGAGACTG 420
QY 421 ATNCCNTNMNGNACNGARGCNGARGARGAATYTTACMNSNGTNGTMSNARGATN 480
DB 421 ATACCTTTGAGAGTCACTGAAGAGCGGAGAGAACTTCACTGCTGCTGCTGCTGCTG 480
QY 481 ATGATGTAATNTYNTNGTNTTYTNAACNTNTGTYNTNTATNGARATGATNTATYTG 540
DB 481 ATGATGTAATCTCTTCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
QY 541 TAYWGNARGTMSNARGCNGARGCNGCNGCARGAAYGCMNSNGATYATYTMNGN 600
DB 541 TACGAGAAAGTCTTAAGCGGAGAGAGCGACAGCAAGAAAATGCTGCTGACTACCTGCT 600
QY 601 ATNCCMNSNGARAAYARGARAAYMSNGCNGTNGTNGARGAR 645
DB 601 ATCCCTTCAAGAGAACAGAGAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 645

RESULT 7
AAC67836
ID AAC67836 standard; cDNA, 2220 BP.
XX

| | |
|----|---|
| XX | AA067836; |
| XX | 15-FEB-2001 (first entry) |
| DE | Rat beta3 subunit cDNA. |
| XX | |
| XX | Rat: beta sub-unit; beta3; analgesic; anticonvulsant; cerebroprotective; |
| KW | vasotropic; cardiac; nootropic; cytostatic; dermatological; |
| KM | gene therapy; voltage-gated sodium channel; pain; epilepsy; stroke; |
| KW | ischemia; heart disease; Jacobsen Syndrome; |
| KW | familial nonchromaffin paraganglioma; phenyleketonuria; |
| XX | Charcot Marie Tooth disease; ss. |
| OS | Rattus sp. |
| XX | |
| XX | WO20063367-A1. |
| XX | |
| PD | 26-OCT-2000. |
| XX | |
| PF | 24-FEB-2000; 2000WO-EP001783. |
| XX | |
| PR | 15-APR-1999; 99US-0129473P. |
| XX | |
| XX | (WARN) WARNER LAMBERT CO. |
| PA | (UYCA-) UNIV CAMERIDE TECH SERVICES LTD. |
| XX | |
| PI | Cox P, Dixon A, Jackson A, Morgan K; |
| DR | WPI; 2000-665241/64. |
| XX | P-PBDB; AAB36001. |
| PT | Novel nucleic acids encoding a beta-3 subunit from a voltage-gated sodium |
| PT | channel, and their corresponding polypeptides, useful for detecting and |
| PT | treating sodium channel-associated conditions, e.g. pain, epilepsy and |
| PT | stroke. |
| XX | |
| XX | Claim 6, Page 69-70; 88pp; English. |
| XX | |
| XX | The present sequence is given in the claims of a specification relating |
| CC | to a novel family of beta sub-unit proteins from a voltage-gated sodium |
| CC | channel. Human and rat beta sub-units, which have been collectively |
| CC | identified as beta3, have been isolated. The polynucleotides and |
| CC | polypeptides are useful for screening for agonists and antagonists of |
| CC | sodium channels. The agonists, antagonists, proteins and nucleic acids |
| CC | may be used diagnosing of treating diseases or conditions associated with |
| CC | voltage-gated sodium channels, e.g. pain, epilepsy, stroke, ischaemia, |
| CC | heart disease, Jacobsen Syndrome, Familial Nonchromaffin Paraganglioma, |
| CC | Phenyleketonuria and Charcot Marie Tooth disease |
| XX | |
| XX | Sequence 2220 BP; 573 A; 557 C; 561 G; 529 T; 0 U; 0 Other; |

| | | | | |
|-----------------------|------------------|--|----------|-------------|
| Query Match | 71.8% | Score 462.8 | DB 3 | Length 2220 |
| Best Local Similarity | 58.4% | Pred. 1.8e-116 | | |
| Matches 377 | Conservative 145 | Mismatches 123 | Indels 0 | Gaps 0 |
| Qy | 1 | ATGCCGNGNTTAAAYMGNVTTTTCNCNTYNGCNSMNYNTGNTVYATATAYTGGTGNMSN | 60 | |
| Db | 363 | ATGCGTGCTTCAACAGATGTGCTTCCCGTAGCTTCTTAGTGTCTATGTACTAGGGTCAGA | 422 | |
| Qy | 61 | GTNTGTTTCCNGINTGTGTNGARGTNCCMSNGAPACNGARGCNGTNCARSGNAAAYCN | 120 | |
| Db | 423 | GTCTGCTTCCCTGTGTGTGTGGAAATGCCCTCGGAGACAGAAACCGGTGACGGGCAATCCC | 482 | |
| Qy | 121 | ATGARAYTNMGTYATYTNMSNTGTATYTGAAABMGNGARGRTNGARGCNACNACNENGTN | 180 | |
| Db | 483 | ATGAAGCTGAGGCTGATCTCTCTGCTATGAAGAGGAGAGAGGTGAGAGCCACCACTGTGGTG | 542 | |
| Qy | 181 | GARTGGTTTAYMGNCCNGARGSGNNGNARGA YTTTYTATATAYGARTAYMGNAAAYGN | 240 | |
| Db | 543 | GAGTGGTTCTACAGGCGCTGAGGGCGGTAAAGA TTTCTTATATATAGATATGGAATGGC | 602 | |
| Qy | 241 | CAYCARGAGTNGARWNSNCNTTTCARGGNMGNVYNCARTGAAYGKMSNABARAYTN | 300 | |

| | | | |
|----|-----|--|-----|
| Db | 603 | CACACAGGAAGTGGAGAGCCCTTCCAAAGCCGTCGACAGTGGAAATGGAGCAAAGACTG | 662 |
| Qy | 301 | CARGAYGTWMSNATNACNGTNYTNAAYGTNACNTYNAAYGATWSNGNTNTAAYCNTGY | 360 |
| Db | 663 | CAGGACGATCATCATCATCTGTACTCAATGTCACTTGGAAATGACTCTGGCCCTCTACACATGC | 722 |
| Qy | 361 | AAYGNMSNMNGARTTYGARTTYGABRCONCAVYMNCCNTTYGTNAAARACNA CMNGNTN | 420 |
| Db | 723 | AATGTGTCACAGGAGGTTGAAATTCAGAGCACACAGGCCCTTYYGTGAAGCCACGACACTG | 782 |
| Qy | 421 | ATNCCNTYTMNGCNVACNGARGARCGNGNGARGAYTYYACNWSNGTNTNMSNGARATN | 480 |
| Db | 783 | ATACCTTTGCCAGTACTGTAAGAGGCGGAGAAAGACTTCACTCCGTGGTCTTCGGAAATC | 842 |
| Qy | 481 | ATGATGTAAATNYTNYTNGTNTYYTNA CNYNTGTYTNYTNAAGARAATGATNTATYGY | 540 |
| Db | 843 | ATGATGTATCATCTCTCTGCTCTTCTCTCA CCTGTGGCTGTTATATGAGATGATCTATTGC | 902 |
| Qy | 541 | TATWGNAAARGTNSNAAANGCNGARGARCGNGCNCAARGAAYGCMWSNGATYATYYTNGCN | 600 |
| Db | 903 | TACAGAAAGGCTCTTAAAGGCCGAAGAGGCACGACAGGAAAAAGCGCTCTGACTACCTTGCT | 962 |
| Qy | 601 | ATNCCMSNGARAAYAARGARAAYMSNCGNTNCCNGTNGARGAR | 645 |
| Db | 963 | ATCCTTTCAGAGAACMAAGAGATCTGTGTGTA CTTGTGAGAGAA 1007 | |

| | |
|----|--|
| XX | RESULT 8 |
| XX | AAC90602 |
| XX | ID AAC90602 standard; DNA; 2632 BP. |
| XX | AC AAC90602; |
| XX | PT 13-MAR-2001 (first entry) |
| XX | DE Rat sodium channel beta3 protein Alrx294h5 related sequence. |
| XX | XX Rat; sodium channel beta3 protein; Alrx294h5; pain; sleep disorder; |
| KW | neurodegenerative disorder; mood disorder; muscle contraction; ds. |
| XX | XX Rattus sp. |
| XX | XX WO200069912-A1. |
| XX | XX 23-NOV-2000. |
| XX | PD 12-MAY-2000; 2000WO-US013144. |
| XX | PE 14-MAY-1999; 99US-0134198P. |
| XX | PR (MILL-) MILLENNIUM PHARM INC. |
| XX | PA Curtis RAJ; |
| PI | WPI; 2001-122743/13. |
| DR | XX New rat sodium channel beta-3 subunit gene isolated from a rat dorsal |
| XX | PT root ganglion cDNA library for use in chromosome mapping, forensic |
| PT | medicine, monitoring clinical trials and therapeutics. |
| XX | XX Claim 1; Page 41-42; 145pp; English. |
| XX | XX The present invention provides the protein and coding sequences of the |
| XX | XX rat sodium channel beta3 protein, designated Alrx294h5. This protein is |
| CC | involved in the generation of pain and other sensory or perceptive nerve |
| CC | impulses, in the establishment and endurance of mood, neurodegenerative |
| CC | and sleep disorders, and in the control of muscle contraction, including |
| CC | movements such as the heartbeat, digestion and vascular tone. The |
| CC | sequences can be used in predictive medicine, screening and diagnostic |
| XX | assays, and in pharmacogenomics |
| XX | Sequence 2632 BP; 693 A; 636 C; 611 G; 692 T; 0 U; 0 Other; |

| | | |
|-----------------------|--|--------------------------|
| ID | AAK52345 | standard; cDNA; 1045 BP. |
| AC | AAK52345; | |
| DT | 06-NOV-2001 | (first entry) |
| XX | | |
| DE | Human polynucleotide SEQ ID NO 890. | |
| XX | | |
| KW | Human; cytokine; cell proliferation; cell differentiation; gene therapy; vaccine; peptide therapy; stem cell growth factor; haematopoiesis; tissue growth factor; immunomodulatory; cancer; leukaemia; nervous system disorder; arthritis; inflammation; ss. | |
| OS | Homo sapiens. | |
| PN | WO200157190-A2. | |
| XX | | |
| PD | 09-AUG-2001. | |
| XX | | |
| PF | 05-FEB-2001; 2001WO-US004098. | |
| XX | | |
| PR | 03-FEB-2000; 2000US-00496914. | |
| PR | 27-APR-2000; 2000US-00560875. | |
| PR | 20-JUN-2000; 2000US-00598075. | |
| PR | 19-JUL-2000; 2000US-00620325. | |
| PR | 01-SEP-2000; 2000US-00654936. | |
| PR | 15-SEP-2000; 2000US-00663561. | |
| PR | 20-OCT-2000; 2000US-00693325. | |
| PR | 30-NOV-2000; 2000US-00728422. | |
| XX | | |
| PA | (HYSE-) HYSEQ INC. | |
| PI | Tang YT, Liu C, Dermanac RT, Asundi V, Zhou P, Xu C, Cao Y; | |
| PI | Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW; | |
| PI | Xue AJ, Yang Y, Wejhtman T, Goodrich R; | |
| XX | | |
| DR | WPI; 2001-476283/51. | |
| DR | P-RSDB; AAM79212. | |
| PT | Nucleic acids encoding polypeptides with cytokine-like activities, useful in diagnosis and gene therapy. | |
| XX | | |
| PS | Claim 1; Page 2934-2935; 6221dp; English. | |
| CC | The invention relates to polynucleotides (AAK51456-AAK53435) and the encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to cytokine, cell proliferation or cell differentiation or which may induce production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities, e.g. stem cell growth factor activity, haematopoiesis regulating activity, tissue growth factor activity, immunomodulatory activity and activin/inhibin activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the sequence listing were missing at the time of publication | |
| SQ | Sequence 1045 BP; 222 A; 289 C; 314 G; 220 T; 0 U; 0 Other; | |
| Query Match | 65.9%; Score 424.8; DB 4; Length 1045; | |
| Best Local Similarity | 58.7%; Pred.No.3.2e-106; | |
| Matches | 348; Conservative 131; Mismatches 114; Indels 0; Gaps 0; | |
| OY | 1 ATGCCGNCNTTAAAGMGNVTTCCTCCTNGCNMSNATNGTNVTAINTAATTGGGTNNNS 60 | |
| Ddb | 302 ATGCCTGCCTCAATAAGATTGTTCCCTCGCTTCTCTCGTGCTATCTACTGAGGTCAGT 361 | |
| OY | 61 GTTNGYTTCGNCNNTGTGTGARGBTNCNMSGARACNGARCGNATCARGNAAVCNN 120 | |
| Ddb | 362 GTCTGCTTCCCTGTGTGTGAATGTGCTCTCGAGACGAGGCCGTGACAGGCAACC 421 | |

| | | | |
|----|-----|---|-----|
| QY | 121 | ATGAARNTMNGTGTATATMSNTGTATGAABMGNGARGARGTNGAGCNACNMGNTGN | 180 |
| Db | 422 | ATGAAGCTGCCGTGTGATCTCTCTCGATGANGAGAGAGAGGTGAGGCCAACACGGTGGTG | 481 |
| QY | 181 | GARTGGTYYTAYMGNCCNGARGGNGNARGA YTTYYTNATNTAYGARTAYMGNAAVGN | 240 |
| Db | 482 | GAATGGTTCACAGGCCCGGAGGCGGTAAAGATTTCTTAATTTACGAGTATGCGAATGGC | 541 |
| QY | 241 | CAYCARGARGTNGAAMSNCCNTTTCARBGNMNTYNTCATGTGAAVYGNMNSNARGA YYTN | 300 |
| Db | 542 | CACCGAGAGGTGGAGAGGCCCTTTTACGGGGCCCTCGCAGTGTGAATGGCAGCAAGGACTTG | 601 |
| QY | 301 | CARGAYGNMNSNATYACNGTNTYNTAAVGTNACNTYNTAAVGA YWSNGNTNTNTAYACNTGY | 360 |
| Db | 602 | CAGACGCTGTCCATCACTGTGCTCAACGTCACTCTGAACGACTCTGGCCTCTACACTTG | 661 |
| QY | 361 | AAVGTNMSMNGNARTTYGARTTYGARGCNCA YMGNCNTTTCYTNAPARACNACMNGNTN | 420 |
| Db | 662 | AATGTGTCCCGGAGTGTGAGTTTGAAGGCGCACTCGGCCCTTTGTGTGAAGACGACGCCGCTG | 721 |
| QY | 421 | ATNCCNNTNMNGTNAACNGARBARCNGNNGARGATYYTACMMSNGTNTGNMNSNGARATN | 480 |
| Db | 722 | ATCCCCCTPAAGAGTACCGAGGAGGCTGGAGAGACCTTCACTGTGGTCTTAGAAATC | 781 |
| QY | 481 | ATGATGTATATYNTYNTYNTGNTYYTNTACNTYNTGGYNTYNTATNGARATGATNTATYTG | 540 |
| Db | 782 | ATGATGTACATCTCTTGTGTCTTCTCAACCTTGTGGCTGCTCATCGAGATGATATATTTGC | 841 |
| QY | 541 | TATMGNAARGTNMSNARCGNARGARGRCNGNCNARGAARAYTCMMSNNGATYA | 593 |
| Db | 842 | TACGAAAGGTCTCAAAAAGCCGAAAGGCGCAAGGCAAGAAAAGCGCTTAAATGTCCA | 894 |

RESULT 11
 AAH98320
 ID AAH98320 standard; cDNA; 978 BP.
 XX
 AC AAH98320;
 XX
 DT 12-OCT-2001 (first entry)
 XX
 DE Human EST-derived coding sequence SEQ ID NO: 177.
 XX
 KW Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
 KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
 KW diagnostics; forensic test; gene mapping; genetic disorder; biodiver
 KW gene therapy; nutrition; ss.
 XX
 OS Homo sapiens.
 XX
 PN W0200154477-A2.
 XX
 PD 02-AUG-2001.
 XX
 PF 25-JAN-2001; 2001MO-US002687.
 XX
 PR 25-JAN-2000; 2000US-00491404.
 PR 17-JUL-2000; 2000US-00617746.
 PR 03-AUG-2000; 2000US-00631451.
 PR 15-SEP-2000; 2000US-00663870.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YF, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
 PI Cao Y, Dmanac RA, Zhang J, Werhman T;
 XX
 DR WPI; 2001-476164/51.
 DR P-PSDB; AAM23661.
 XX
 PT Isolated polypeptide for treatment of diseases, diagnostics, raising
 PT antibodies and research use.
 XX
 PS Claim 1; Page 299-300; 1275pp; English.

XX The present invention provides the protein and coding sequences of novel
 CC proteins from a variety of organisms, including human, dog, cat, horse,
 CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
 CC urchin and tomato. These were derived from expressed sequence tags (ESTs)
 CC from the organism of interest. They can be used in diagnostics,
 CC forensics, gene mapping, identification of mutations, to assess
 CC biodiversity and for nutritional purposes. The present sequence is a cDNA
 CC of the invention

XX Sequence 978 BP, 193 A; 283 C; 293 G; 209 T; 0 U; 0 Other;

XX Query Match 63.9%; Score 412.4; DB 4; Length 978;

XX Best Local Similarity 56.3%; Pred. No. 8.1e-103;

XX Matches 363; Conservative 139; Mismatches 141; Indels 2; Gaps 2;

Qy 1 ATGCCGNCNTTYAAYMGNNTTYCCNTYNGCMNSNTYNTNATNTATYATGGTMSN 60
 Db 295 ATGCTGCTCCATTAAGATTGTTCCCTGCTCTCTCGTCTTACTTACTGGGTACGT 354
 Qy 61 GTNTGTYTTCNGNTGTYGTNGARGTNCNMSNGARCGNCTNCARGNAAVCCN 120
 Db 355 GTCTGCTTCCCTGTGTGTGTGAAGTGCCCTCGAGACGAGCGCTGACGCAACCC 414
 Qy 121 ATGAARYTNMNTGYATNTMSNTGYATGAABMGNGARGTNGARCGNACNACNGTNTN 180
 Db 415 ATGAAGCTGGCGCTCATCTCTCGATGAAGAGAGAGAGGAGGCGCACACGCTGTG 474
 Qy 181 GARTGCTTYTAYMGNCCNGARCGNAGNAAAGATTTTNTATNTAYGARTAYMGNAAVGN 240
 Db 475 GAATGCTTCTACAGCCCGAGGGGGGTAAAGATTCTTATTCAGATTCGGAATGCG 534
 Qy 241 CAYCARGARGTNGARMSNCCNTTYCARGNMNGNTNCACTGAAVYGNMSNAAAGAYTN 300
 Db 535 CACGAGAGGTGAGAGCGCCCTTACAGGCGCGCTGACGTGAATGCGACGACGACCTG 594
 Qy 301 CARGAYTNMNSNATNACNGTNTYTAAYGTNAONTNAYGMSNGNTYNTAYACNTGY 360
 Db 595 CAGAGCGTGTCACTCATCTGCTCAACGCTCACTGAAACGACTGCGCTCTACACCTGC 654
 Qy 361 AAYGTNMSNMGARTTYGARTTYGARGCNCAMNCCNTTYTNAARACNACMNGYTN 420
 Db 655 AATGTGTCGCCGAGATTGAGATTGAGCGCATCGGCCCTTGTGAAGACACCGCGCTG 714
 Qy 421 ATNCCNTYTNMNGTNAACNGARCGNCGNAGARGAAYTTTACNMSNGTNGMSNARATN 480
 Db 715 ATCCCCCTAAGAGACACCGAGAGGCTGGAAGACCTTCACTCTGTGTCTCAGAAATC 774
 Qy 481 ATGATGTAYATNTYNTNNTYNTYTAACNTNTGTYNTNTATNTAGAAATATATYTG- 539
 Db 775 ATGATGTACATCTTCTGTGCTTCTCTCACTTGTGTGCTCATCGAGATATATATTCG 834
 Qy 540 YTAAGNARGTNMSNARCGNCGNAGARCGNCGNAGARGAAYGCMNSNGATY-YYTN 598
 Db 835 CTACAGAACGCTGATTCACAGACGAAACGAGCCGCCCAAGAAACCGATGTGCTACCTTG 894
 Qy 599 CNAATCCNMSNAGAPAAVYAGARAAVYMSNGNTNCCNGTNGARG 643
 Db 895 CGATTCCATTGTGAAGACGAGAAATCTTCTGTGCTGCGGGG 939

RESULT 12

ADSI1487 ID ADSI1487 standard; DNA; 978 BP.

XX AC ADSI1487;

XX DT 16-DEC-2004 (first entry)

XX DE Human therapeutic contig DNA - SEQ ID 1724.

XX KW antiinflammatory; neuroprotective; antihaemic; cytostatic; vulnerary;
 KW inflammatory; haematopoiesis; immunity; neurodegenerative; stem cell;

KW aplastic anaemia; cancer; wound healing; gene therapy; ds; gene.

XX Homo_sapiens.

XX MO2004080148-A2.

XX 23-SEP-2004.

XX 30-SEP-2003; 2003WO-US030720.

XX 02-OCT-2002; 2002US-0416186P.

XX (NUVE-) NUVELO INC.

XX Tang YT, Asundi V, Ren F, Zhang J, Wehrman T, Wang Z, Ma Y,
 PI Wang D, Chen R, Zhao QA, Wang J, Ghosh M, Xue AJ, Weng G, Zhou P;

XX WPI: 2004-668857/65.

XX P-PSDB; ADSI2085.

PT New polynucleotide, useful in preparing a composition for diagnosing or
 PT treating inflammatory, neurodegenerative or stem cell disorders, e.g.,
 PT aplastic anemia or cancer for promoting wound healing.

XX Example 2; SEQ ID NO 1724; 718bp; English.

CC The invention relates to a novel isolated polynucleotide and the encoded
 CC polypeptide. The molecules of the invention demonstrate antiinflammatory,
 CC neuroprotective, antihaemic, cytostatic and vulnerary activities and may
 CC be useful in preparing a composition for diagnosing or treating
 CC inflammatory, haematopoietic, immune, neurodegenerative or stem cell
 CC disorders, such as aplastic anaemia or cancer, as well as for promoting
 CC wound healing. The molecules may also be utilised during gene therapy
 CC procedures. The current sequence is that of a human therapeutic contig
 CC DNA of the invention. The current sequence is not shown explicitly within
 CC the specification but can be accessed from the WIPO web-site.

XX Sequence 978 BP, 193 A; 283 C; 293 G; 209 T; 0 U; 0 Other;

XX Query Match 63.9%; Score 412.4; DB 13; Length 978;

XX Best Local Similarity 56.3%; Pred. No. 8.1e-103;

XX Matches 363; Conservative 139; Mismatches 141; Indels 2; Gaps 2;

Qy 1 ATGCCGNCNTTYAAYMGNNTTYCCNTYNGCMNSNTYNTNATNTATYATGGTMSN 60
 Db 295 ATGCTGCTCCATTAAGATTGTTCCCTGCTCTCTCGTCTTACTTACTGGGTACGT 354
 Qy 61 GTNTGTYTTCNGNTGTYGTNGARGTNCNMSNGARCGNCTNCARGNAAVCCN 120
 Db 355 GTCTGCTTCCCTGTGTGTGTGAAGTGCCCTCGAGACGAGGCGCTGACGCAACCC 414
 Qy 121 ATGAARYTNMNTGYATNTMSNTGYATGAABMGNGARGTNGARCGNACNACNGTNTN 180
 Db 415 ATGAAGCTGGCGCTCATCTCTCGATGAAGAGAGAGAGGAGGCGCACACGCTGTG 474
 Qy 181 GARTGCTTYTAYMGNCCNGARCGNAGNAAAGATTTTNTATNTAYGARTAYMGNAAVGN 240
 Db 475 GAATGCTTCTACAGCCCGAGGGGGGTAAAGATTCTTATTCAGATTCGGAATGCG 534
 Qy 241 CAYCARGARGTNGARMSNCCNTTYCARGNMNGNTNCACTGAAVYGNMSNAAAGAYTN 300
 Db 535 CACGAGAGGTGAGAGCGCCCTTACAGGCGCGCTGACGTGAATGCGACGACGACCTG 594
 Qy 301 CARGAYTNMNSNATNACNGTNTYTAAYGTNAONTNAYGMSNGNTYNTAYACNTGY 360
 Db 595 CAGAGCGTGTCACTCATCTGCTCAACGCTCACTGAAACGACTGCGCTCTACACCTGC 654
 Qy 361 AAYGTNMSNMGARTTYGARTTYGARGCNCAMNCCNTTYTNAARACNACMNGYTN 420
 Db 655 AATGTGTCGCCGAGATTGAGATTGAGCGCATCGGCCCTTGTGAAGACACCGCGCTG 714
 Qy 421 ATNCCNTYTNMNGTNAACNGARCGNCGNAGARGAAYTTTACNMSNGTNGMSNARATN 480

QY 301 CARGAYGTMNSNATNACNGTNTNAAYGTNACNYTNAAYGAYWNSNGNTNTAYACNTGY 360
 Db 328 CAGGACGTGTCCATCACTGTGTCAAGTCACTGTGACACTCTGACCTCTACACTGC 387
 QY 361 AAYGTNMSNMNGARTTYGARITTYGARGCNCAVMGNCNTTYGTNAARACNACMGNTYN 420
 Db 388 AATGTGTCCCGGAGTTTGAGTTTGAGGCGCATCGGCCCTTGTGAAGACGACGCGGCTG 447
 QY 421 ATNCCNYTNMGNT 434
 Db 448 ATCCCCCTAAGAT 461

Search completed: April 7, 2005, 21:46:50
 Job time : 438 secs

TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 13053
LENGTH: 30337
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc feature
LOCATION: (1) - (30337)
OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13053

Query Match 9.0%; Score 57.8; DB 4; Length 30337;
Best Local Similarity 44.3%; Pred. No. 2.3e-05;
Matches 62; Conservative 23; Mismatches 55; Indels 0; Gaps 0;

Qy 442 GARGCNGNGARGAYTTACWMSNGTGNMNGARATNATGATGATAYATNTYNTGNTN 501
Db 608 GCACCCACAGACAGATGCATCATCGTCTGATGATGATGATGATGATGATGATGATG 667
Qy 502 TTYTNAACNTYNTGNTYNTATGATGATGATGATGATGATGATGATGATGATGATGAT 561
Db 668 GTGTGACCAATATGCTGCTGCGACAGATGATGATGATGATGATGATGATGATGATG 727
Qy 562 GARGAGCNGCNCARGAAA 581
Db 728 ACGGAGACTGCTGCACAGCA 747

RESULT 5
US-09-806-708B-22
Sequence 22, Application US/09806708B
Patent No. 6784342
GENERAL INFORMATION:
APPLICANT: The University of British Columbia
TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
FILE REFERENCE: 4810-58741
CURRENT APPLICATION NUMBER: US/09/806,708B
CURRENT FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: US 60/147,133
PRIOR FILING DATE: 1999-08-04
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn version 3.0
SEQ ID NO 22
LENGTH: 1141
TYPE: DNA
ORGANISM: Artificial sequence
FEATURE:
NAME/KEY: promoter
LOCATION: (1) - (1141)
OTHER INFORMATION: consensus sequence of A.T., L.A., and B.N. PAB1 promoters
US-09-806-708B-22

Query Match 8.7%; Score 56.4; DB 4; Length 1141;
Best Local Similarity 11.2%; Pred. No. 3.1e-06;
Matches 70; Conservative 232; Mismatches 323; Indels 2; Gaps 1;

Qy 8 CNTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 67
Db 134 BDTVDHVVTAANNAATTCMDXDDKTRTWWKNNATGMDDTTKHMMNNNGCBTGT 193
Qy 68 TYCCGNTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 127

Db 194 WMAVYKTDKDBSKBGMVYGMWKMNSYDVTYVWVDDKCKRYRBRWVTRGMRMYV 253
Qy 128 TTMGNTGATATWNSNTGYATGAARMNGARGATNGARGCNAACNGTNGARGGT 187
Db 254 AMBTAFHRRRYNNGWTBAMAYRBTNNNNNNNAKAMCRKAKYGNRABVNSTCTTWSKT 313
Qy 188 TTYTAMGNCNGARGNGGNAARAGATYTTTATNTATYGAATYMGAAAGNCAATCARG 247
Db 314 TKVTSWANNCRDGNKDKHKKWMSAAGVYNNNNNNNNWYTKARHBAARMWVHSA 373
Qy 248 ARGNGARMSNCNTTYCARGMNGYNTNCAATGAAAYGGWMSAARAGAYTNCARGAYG 307
Db 374 WKKHAAAHYSRKQWBYRKTYVNNNNGTTHMKRMAWYTKMDMBGTYYNNNNNGR 433
Qy 308 TMSNATNACNGTNTYNAAYGTNACNTYNAAYGAYWNGNGYNTTAYACNTGYATGNTW 367
Db 434 TYVGMTKMKWMTYTKWANNCKRAMDHTCTHNTTMMWTKYNNNCKYWSMTNGSKH 493
Qy 368 SMMGNGARTTYGATTTYGAGCNCAYMGNCCNTTYGTAARACACMNGYNTNATNCNTY 427
Db 494 RBAAAYTYWMMWRRYAHANNNNWDYWKACWYKTBVCSKWNNTYAAWTKSSMNTS 553
Qy 428 TTMGNTNACGARGCNGCNGARGAYTTACWMSNGTGNMNGARATNATGATGAT 487
Db 554 RYRMTKTNMSRMSDTRSMGRANNYARABHYGYKNTTRWBSHTWBHBAAGAAHYWMB 613
Qy 614 MMYBAKCHCMKAWYKAKYAGAGSNNNNNNNNNNNNNNNNNNATCARDYAAASHWYMAN 673
Qy 548 ARGTMNSNA--RCNCRARGCNGCNCARGAAAAGCMWSNCAATYATYNTGNCATNCC 605
Db 674 AKWYTYBAANNAATYTAANWGWCMNATDTRFRYKNNNNNNNAGTWKNNNNNAKQASA 733
Qy 606 NWSNGARAAYARGARAAYWSNGCNGT 632
Db 734 AKVYAAAAYAKKGMWMAKMAARGW 760

RESULT 6
US-08-232-463-14/C
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768

Db 241 TCTYKSYTCKRSYTTTSTSKGKMGCTKSRMSYTMMSKSYTTCGSKKKRMYMSAGA 300
Qy 397 CCNTTGTNARACNACNMGNTNATNCNTNMGNGTNAACGARGCGNGGARGAY 456
Db 301 WYAMWSMWCAMCMAMGMRGAMKSRKRYTMAKSCMYCAMKSCSARAKGRCCT 360
Qy 457 TTYACNMSGNGTNGTNGTNGATATATGATGATATNTNTNTNTNTNTNTNTG 516
Db 361 TTKYKMTTTCYCMYKRCMWSMCAVCTGYCTYTKTWTGCTWACTGYTYSK 420
Qy 517 YTNNTNAT 524
Db 421 TKMYTMT 428

RESULT 9
US-09-621-976-18033
Sequence 18033, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 18033
LENGTH: 474
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 16
OTHER INFORMATION: n=a, g, c or t
US-09-621-976-18033

Query Match 7.8%; Score 50.4; DB 4; Length 474;
Best Local Similarity 8.7%; Pred. No. 9,4e-05;
Matches 29; Conservative 142; Mismatches 164; Indels 0; Gaps 0;
Qy 210 RGATTTTATNTATYAGRTYMGNAAYGNCAYCARGATNGRNSCCTTYCARGG 269
Db 14 SGCYCCMAKMSYKGRMYKSMRGSSCYKSCMCMKSCYSKSYKTTTWTWAWM 73
Qy 270 NMGNVTCARTGAAYGSMGNAARAYTNCARGAVGTMSNATNACNGTNTNAAVGT 329
Db 74 WTTTGGKMAARRRSGGKTTYMCKSKKTKSCMAGRWKGKTYISRMYYCYKACATYMW 133
Qy 330 NACNYTNAAYGAYNSNGNTNTAYACNTGYAAYGTMSNMGNGARTTYGARTYGARGC 389
Db 134 KRWYSSCCMWTYTGCGSMWTTTMMRRKSKYKRWTKGKKKTTMMAMCVTTWRSY 193
Qy 390 NCAYMGCCNTTGTNARACNACNMGNTNATNCNTNMGNGTNAACGARGCGNGG 449
Db 194 WMMMRRAAAKTYCYCMKSKTMCMMCCMCCRARSCCMSKMSYTMCCYTYM 253
Qy 450 NGARGATTYACNMSGNGTNGTNGTNGATATATGATGATATNTNTNTNTNTNTAC 509
Db 254 MYKGRMYTMMRGMMKRMKMYTKKKSKMSKMSKMRMMAKTTYTMMYYTYYKRM 313
Qy 510 NYTNTGTYTNTATNGARATGATNTAYTAYTAYM 544
Db 314 CCYMRKTTYCMWMMYSRWMSMTARGAWMWMCY 348

RESULT 10
US-09-621-976-2813/C
Sequence 2813, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:

APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 2813
LENGTH: 832
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: 235..399
US-09-621-976-2813

Query Match 7.8%; Score 50; DB 4; Length 832;
Best Local Similarity 9.4%; Pred. No. 0.00021;
Matches 34; Conservative 157; Mismatches 169; Indels 2; Gaps 1;
Qy 2 TGCCNGCTTYAAYMGNYTNTTYCCNTNGCNSNGTNTNTATNTATYTGCGTMSNG 61
Db 365 TGTYYTWTWKTYWYTYTTRMMWTKKARRYWTKSYTACASRYKTYGWWYTMK 306
Qy 62 TNGYTYCCNGTNTGYGTNGARGTNCNMSGNARGCNGTNCARGNAAYCCNA 121
Db 305 RMSTRWYCYWCKCVCYRGRCAWYTMARGMWSYAMGKWSRMSMCTRYKKS 246
Qy 122 TGAARYTNGTGYATWMSYTGATGAABMGNGARGRTNGARGCNAACNCGTNGTNG 181
Db 245 TYWTMKTCTATWCYWKYKMSKTCWGSRGTYSTSTSYMYM--SMWTMC 188
Qy 182 ARTGTYTYVMGNCNGARGGNGNARGAYTYTNTATNTAYGARTAYGNAAYGNGC 241
Db 187 WMMGRWSTYTYMAGKWKWRVATWRRAMWMMWMAATWMMWMAACMSRGAAYRT 128
Qy 242 AYCARGARTNGARNSCCTTYCARGANNGTNTNCAFTGGAAYGMSNARGAYTNC 301
Db 127 MMMGYWYWRKSKYRTRCAMAAYAMTKRSYWCWPKWYRCMMWMAAYGKTYMMR 68
Qy 302 ARGAYGTMSNATNACNGTNTNAAVGTNACNTNAAVGYNSGNYTNTAYACNTGYA 361
Db 67 ACWTRTWMMWAMWMMWMTMMWYWPAMKRRMMKMSWSMMWAMWNTWPAAR 8
Qy 362 AY 363
Db 7 MW 6

RESULT 11
US-09-621-976-15639/C
Sequence 15639, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 15639
LENGTH: 505
TYPE: DNA
ORGANISM: Homo sapiens
US-09-621-976-15639
Query Match 7.5%; Score 48.6; DB 4; Length 505;
Best Local Similarity 11.1%; Pred. No. 0.00035;

| | Matches | 41; | Conservative | 140; | Mismatches | 187; | Indels | 0; | Gaps | 0; |
|----|---------|---|--------------|------|------------|------|--------|----|------|----|
| Qy | 244 | CARGARGINGARMSNCNTTYTCARGGMNGNTNCARTGGAAYGVGWSMAARGAAYTTCNAR | 303 | | | | | | | |
| Db | 371 | MAAAKMRMAAAGGSYCGMTSTYSSKATGRGSGATKKRPMWTSYGGWWTISKYCKTKGK | 312 | | | | | | | |
| Qy | 304 | GAYGTMSNATNACNGTNYTNAAYGTNACNYTNAAYGAWSNAGSYVTYTAAYCACTGAAY | 363 | | | | | | | |
| Db | 311 | YTGWKSXKTRWTCISWRKYTMMMSGCWARSMSWARSYTSMMACCMMSASAYRARBMY | 252 | | | | | | | |
| Qy | 364 | GTNMSNMGNAGRTTYGARFTTYGARGCNCAYMGNCCNTTYGTNMAAPACNACMNGYTNATN | 423 | | | | | | | |
| Db | 251 | GARSMRPAAGAWMPARRRGKTKARBSKMSMSKMSMSMSAGKAMCRRMMMSCRMGSYS | 192 | | | | | | | |
| Qy | 424 | CCNTNMNGTNAACNGARGCNGCNGNARGAITYTACNWSGNTGTMSNCAARTNATG | 483 | | | | | | | |
| Db | 191 | CMGSCMSCRGTCAAMWYARAPKRYASMSGKTYMGCWCTAKCARMGYRBSRSTGS | 132 | | | | | | | |
| Qy | 484 | ATGTAYATNNTYNTGNTTYYTNAACNTNTGTNTYNTNATNGARATGATNTAYGYTAY | 543 | | | | | | | |
| Db | 131 | RGMYKRRRKMYTMMKMYMMSWMCYMGAAWYGSABARMTASMAACMCKCSMMCMMSMS | 72 | | | | | | | |
| Qy | 544 | MGNDAARGTMSNAARCGNGARGCNGCNGNCARGARAAVGCNWSGAYTAYTTCNATN | 603 | | | | | | | |
| Db | 71 | MMMCWBSRYCWMWSGKWYCSOCGYCCSACBMCYCWTRPMKMSWSSRRKSMCCRYSMSMS | 12 | | | | | | | |
| Qy | 604 | CCNWSNGA | 611 | | | | | | | |
| Db | 11 | AYRYSKR | 4 | | | | | | | |

```

RESULT 12
US-09-621-976-8976
: Sequence 8976, Application US/09621976
: Patent No. 6639063
: GENERAL INFORMATION:
: APPLICANT: Dumas Milne Edwards, J.B.
: APPLICANT: Jobert, S.
: APPLICANT: Giordano, J.Y.
: TITLE OF INVENTION: ESTs and Encoded Human Proteins.
: FILE REFERENCE: GENSET.054PR2
: CURRENT APPLICATION NUMBER: US/09/621,976
: CURRENT FILING DATE: 2000-07-21
: NUMBER OF SEQ ID NOS: 19335
: SOFTWARE: Patent.pm
: SEQ ID NO 8976
: LENGTH: 399
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-09-621-976-8976

```

| | | | | |
|-----------------------|--------|-------------------|-----------------|------------------|
| Query Match | 7.2% | Score 46.2; | DB 4, | Length 399; |
| Best Local Similarity | 10.2%; | Pred. No. 0.0015; | | |
| Matches | 32; | Conservative 129; | Mismatches 153; | Indels 0; Gaps 0 |

| | | | | |
|----|-----|-------------------------------------|---------------------------------|-----|
| Qy | 135 | YATWWSNTGYATGARMGNGARGAGTGTGARGCNA | CNACNGTNGTNGARGTGTGTTATWG | 194 |
| | | | | |
| | | ::: ::: | ::: ::: | |
| Db | 5 | YMCWAKKAKMRAMTBRAYWGYITRSGSKSPRAKWS | WMKKRRRRRRRAMWMSMCMKKSK | 64 |
| | | | | |
| | | ::: ::: | ::: ::: | |
| Qy | 195 | NCCNGARGGNGNNAARGAYTTYTYTNTATYATGAP | TAIYMGNAAYGNCAYCARGARGTNGA | 254 |
| | | | | |
| | | ::: ::: | ::: ::: | |
| Db | 65 | SWMSRSGWMTTRMKRGGAASVAGYSSWMTYMTFR | RYRYRKKCATKTPRAAGMCKAOMA | 124 |
| | | | | |
| | | ::: ::: | ::: ::: | |
| Qy | 255 | RMSNCNTTYCARGGNGNYTNCARTGGAAYGNG | SNNAARGYYTNCARGAYGTWWSNAT | 314 |
| | | | | |
| | | ::: ::: | ::: ::: | |
| Db | 125 | WMAIYAKWYTMWRRPAMKTYWAMMSKSR | RRRRRRAIYAWWYTMARPTWMGASCYRGAY | 184 |
| | | | | |
| | | ::: ::: | ::: ::: | |
| Qy | 315 | NACNGTYNTNAAYGTNACNNTYNAAGAYWSNG | NTNTNTAYACNTGYAAAYGTWWSNMGGA | 374 |
| | | | | |
| | | ::: ::: | ::: ::: | |
| Db | 185 | MASAGMYTMMWYTYMWRKCMYTSAGMSW | MRKMTTRCASYSCSSYCMWGAKKMMYKTSRW | 244 |
| | | | | |
| | | ::: ::: | ::: ::: | |
| Qy | 375 | RTTYGARTTYGARGNCAYMGNCCTTYTG | TNAPARCNACNMGNTYNTAACCNTYTMNGGT | 434 |
| | | | | |
| | | ::: ::: | ::: ::: | |

Db 245 SYRYSSYRCKTKYRRSCCSMSCYKTYTSMYCASCTYYSYKTRASCCMMCCMKRRKM 304

QY 435 NACNGARGARGCNG 448
| : : : : | :
Db 305 MAMMYMMKMYCKK 318

```

RESULT 13
US-09-949-016-787
; Sequence 787, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO010107
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 787
; LENGTH: 1342
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-787

```

| Query Match | 7.1% | Score 46 | DB 4 | Length 1342 |
|-----------------------|-----------------|--|----------|-------------|
| Best Local Similarity | 28.4% | Pred. No. 0.0053 | | |
| Matches 137 | Conservative 62 | Mismatches 280 | Indels 3 | Gaps 1 |
| QY | 89 | CNWSNGARACNGARCNCGTNCARGNAAYCCNATGAARYYMNGTGYATWNSNTGYATGA | 148 | |
| Db | 214 | CCCCGGTGCTGGAGAGCTGTATTAATGGGACAGATGCTCGGTAAATGCACTTCTCCACCT | 273 | |
| QY | 149 | ARMNGARGARGTNGARGCNAACNACNCGTNGTNGARGTGTYYTAYMGNCCNARGNGG-- | 206 | |
| Db | 274 | TTGGCCCCGTGGGTGATGCTCTAAACGTGACCTGGAAATTTTCCTCTTACACGGGGGAC | 333 | |
| QY | 207 | -NAARGAYTYYTNATNTAAYGARTAYMGNAAAYGNCAYCARGARGTNGARMSNCNTYYC | 265 | |
| Db | 334 | CTGAGCACTTGTATTTCTACTACACATAGATCCCTTCCAAACCATGATGGGCGGTTTA | 393 | |
| QY | 266 | ARGNMGNTYNTCARTGGAAYGGMNSNAARGAYYTNCAARGAYGTWNSNATNAACNTYNTNA | 325 | |
| Db | 394 | AGAGACGGGGTGCTTTGGGATGGGAATCCTGAGCGGTACGATGGCCCTCATCTTCTCGGA | 453 | |
| QY | 326 | AYGTNACNTYNTAAVGAATWSNGANTNTNAYACNTGYAAVGTWMSMNGARPTTYGARTTYG | 385 | |
| Db | 454 | AACGCACTTCGACGACAAATGGGACATACACCTCCAGGTAAAGAACCCACCTGATGTG | 513 | |
| QY | 386 | ARGNCAYMGNCCNTTYGTNABARACNACMNGYNTNATNCNYYTNGTNGTNAACNARGARG | 445 | |
| Db | 514 | ATGGGGGTGATAGGGGAAGATCCGGCTCAGCGCTGCGACACACTGTACGCTTCTTGAGATCC | 573 | |
| QY | 446 | CNNGNARGAYTTYAACNWSNGTNGTWNNSNGARATNATGATGATAYATNTYNTNGTNTYY | 505 | |
| Db | 574 | ACTTCCTGGGCTCTGGCCATTTGGCTCTGGCTGTGACATGATCATATATAGTAATTTGAG | 633 | |
| QY | 506 | TNACNTYNTGTYNTYNTNATNGARATGATNTATYTGTYTAYMGAAGAATWMSNAARGCNARG | 565 | |
| Db | 634 | TGTCCTCTTCCACGACATTAACGGAAAAAGCATGGGCGGAAAAGAGCTCATTAAGGTGTGG | 693 | |
| QY | 566 | AR 567 | | |
| Db | 694 | AG 695 | | |

| | | | | |
|-----------------------|------------------|-------------------|-----------|--------------|
| Query Match | 7.1% | Score 46; | DB 4; | Length 1708; |
| Best Local Similarity | 28.4%; | Pred. No. 0.0066; | | |
| Matches 137; | Conservative 62; | Mismatches 280; | Indels 3; | Gaps 1 |

RESULT 15

US-09-621-976-8976/c
; Sequence 8976, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.

Query Match 6.9%; Score 44.8; DB 4; Length 399;
Best Local Similarity 7.7%; Pred. No. 0.004;
Matches 28; Conservative 161; Mismatches 172; Indels 2; Gaps 1.

Search completed: April 7, 2005, 23:32:04
Job time : 167 secs

Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2005, 21:39:28 ; Search time 502 Seconds
(without alignments)
7786.259 Million cell updates/sec

Title: US-09-977-579-2-REV
Sequence: 1 ATGCCGNCNTTYAAYMGNNT.....SNGCNGTNCNGTNGARGAR 645

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5615251 seqs, 3030001701 residues

Total number of hits satisfying chosen parameters: 11230502

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA: *
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US10F_NEW_PUB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
22: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | length | DB | ID | Description |
|------------|-------|-------------|--------|----|---------------------|---------------------|
| 1 | 466 | 72.2 | 1261 | 11 | US-09-977-579-4 | Sequence 4, Appl1 |
| 2 | 462.8 | 71.8 | 645 | 13 | US-10-029-191-21 | Sequence 21, Appl |
| 3 | 462.8 | 71.8 | 2220 | 11 | US-09-977-579-3 | Sequence 3, Appl1 |
| 4 | 462.8 | 71.8 | 2632 | 13 | US-10-029-191-22 | Sequence 22, Appl |
| 5 | 462.8 | 71.8 | 3108 | 13 | US-10-029-191-1 | Sequence 1, Appl1 |
| 6 | 158.4 | 24.6 | 1335 | 18 | US-10-723-860-2247 | Sequence 2247, Appl |
| 7 | 158 | 24.5 | 657 | 13 | US-10-029-191-23 | Sequence 23, Appl |
| 8 | 156.8 | 24.3 | 1414 | 18 | US-10-477-272-1 | Sequence 1, Appl |
| 9 | 152.2 | 23.6 | 1490 | 9 | US-09-917-800A-1554 | Sequence 1554, Ap |
| c 10 | 105 | 16.3 | 407 | 17 | US-10-276-774-718 | Sequence 718, Appl |
| 11 | 81.2 | 12.6 | 807 | 17 | US-10-401-916-12 | Sequence 12, Appl |

| | | | | | | |
|----|------|------|------|----|--------------------|-------------------|
| 12 | 81.2 | 12.6 | 974 | 17 | US-10-401-916-13 | Sequence 13, Appl |
| 13 | 57.8 | 9.0 | 3583 | 18 | US-10-723-860-6471 | Sequence 6471, Ap |
| 14 | 52.6 | 8.2 | 6217 | 15 | US-10-311-455-911 | Sequence 911, App |
| 15 | 50 | 7.8 | 671 | 14 | US-10-184-644-346 | Sequence 346, App |
| 16 | 50 | 7.8 | 671 | 14 | US-10-184-634-346 | Sequence 346, App |
| 17 | 49.6 | 7.7 | 2415 | 9 | US-09-978-385-3 | Sequence 3, Appl1 |
| 18 | 49.4 | 7.7 | 633 | 17 | US-10-338-110-121 | Sequence 121, App |
| 19 | 49.4 | 7.7 | 691 | 14 | US-10-123-155-16 | Sequence 16, Appl |
| 20 | 49.4 | 7.7 | 691 | 15 | US-10-146-731-16 | Sequence 16, Appl |
| 21 | 49.4 | 7.7 | 691 | 15 | US-10-140-472-16 | Sequence 16, Appl |
| 22 | 49.4 | 7.7 | 691 | 15 | US-10-141-761-16 | Sequence 16, Appl |
| 23 | 49.4 | 7.7 | 691 | 16 | US-10-142-885-16 | Sequence 16, Appl |
| 24 | 49.4 | 7.7 | 691 | 16 | US-10-158-790-16 | Sequence 16, Appl |
| 25 | 49.4 | 7.7 | 691 | 17 | US-10-137-871-16 | Sequence 16, Appl |
| 26 | 49.4 | 7.7 | 691 | 17 | US-10-140-923-16 | Sequence 16, Appl |
| 27 | 49.4 | 7.7 | 691 | 17 | US-10-141-756-16 | Sequence 16, Appl |
| 28 | 49.4 | 7.7 | 691 | 17 | US-10-141-759-16 | Sequence 16, Appl |
| 29 | 49.4 | 7.7 | 691 | 17 | US-10-140-805-16 | Sequence 16, Appl |
| 30 | 49.4 | 7.7 | 691 | 17 | US-10-140-864-16 | Sequence 16, Appl |
| 31 | 49.4 | 7.7 | 691 | 17 | US-10-142-426-16 | Sequence 16, Appl |
| 32 | 48.4 | 7.5 | 1089 | 10 | US-09-813-153-57 | Sequence 57, Appl |
| 33 | 47.8 | 7.4 | 765 | 14 | US-10-123-155-28 | Sequence 28, Appl |
| 34 | 47.8 | 7.4 | 765 | 15 | US-10-146-731-28 | Sequence 28, Appl |
| 35 | 47.8 | 7.4 | 765 | 15 | US-10-140-472-28 | Sequence 28, Appl |
| 36 | 47.8 | 7.4 | 765 | 15 | US-10-141-751-28 | Sequence 28, Appl |
| 37 | 47.8 | 7.4 | 765 | 16 | US-10-142-885-28 | Sequence 28, Appl |
| 38 | 47.8 | 7.4 | 765 | 16 | US-10-158-790-28 | Sequence 28, Appl |
| 39 | 47.8 | 7.4 | 765 | 17 | US-10-137-871-28 | Sequence 28, Appl |
| 40 | 47.8 | 7.4 | 765 | 17 | US-10-140-923-28 | Sequence 28, Appl |
| 41 | 47.8 | 7.4 | 765 | 17 | US-10-141-756-28 | Sequence 28, Appl |
| 42 | 47.8 | 7.4 | 765 | 17 | US-10-141-759-28 | Sequence 28, Appl |
| 43 | 47.8 | 7.4 | 765 | 17 | US-10-140-805-28 | Sequence 28, Appl |
| 44 | 47.8 | 7.4 | 765 | 17 | US-10-140-864-28 | Sequence 28, Appl |
| 45 | 47.8 | 7.4 | 765 | 17 | US-10-142-426-28 | Sequence 28, Appl |

ALIGNMENTS

RESULT 1
US-09-977-579-4 Application US/09977579
Publication No. US20040248240A1
GENERAL INFORMATION:
APPLICANT: Cambridge University Technical Services
TITLE OF INVENTION: A novel family of beta sub-unit proteins from a voltage gated sodi
TITLE OF INVENTION: channel
FILE REFERENCE: 674558-2001
CURRENT APPLICATION NUMBER: US/09/977, 579
PRIOR FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: PCT/EP00/01783
PRIOR FILING DATE: 2000-02-24
NUMBER OF SEQ ID NOS: 47
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 1261
TYPE: DNA
ORGANISM: Homo sapiens
US-09-977-579-4
Query Match 72.2%; Score 466; DB 11; Length 1261;
Best Local Similarity 58.8%; Pred. No. 1.1e-120;
Matches 379; Conservative 145; Mismatches 121; Indels 0; Gaps 0;
Oy 1 ATGCCGNCNTTYAAYMGNNTTYCCYTCGMSNTNTNTATNTATYTGCTGTSN 60
Db 376 ATGCCGNCNTTYAAYMGNNTTYCCYTCGMSNTNTNTATNTATYTGCTGTSN 60
Oy 61 GTTGTGTTTCGNCNTTYGNGARGTNCNWSNGARGCNGTNGARGGAAAYCCN 120

```

Db 436 GTCGTCTCCCTGTGTGTGGAAGTCCCTCGGAGACGGAGGCGGTGACGGGCAACCCC 495
QY 121 ATGAATYTMNGTGYATNWSNTGYATGAARMNGARGGTNGARGCNAACNCTNGTN 180
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 496 ATGAACCTCGCGCTCATCTCTGCATGAAGAGAGAGAGGTGAGGCCACACGGGTGGT 555
QY 181 GARTGTTTAAVWNCNGARGGNGNAAAGATTTTNTATNTAYGARTAYMGNAAYGN 240
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 556 GAATGTTTCTACAGGCGCCGAGGCGGTAAAGATTTCTTATTTTACGATATCGAATGGC 615
QY 241 CAYCARGGTNGARGNSCCTTYTCARGGNGNTYTCARTGAAYGAGNSNAARGAYTN 300
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 616 CACCAAGAGGTGAGAGCCCTTTCAGGGCGCCTCAGATGGAATGACAGCAAGACCTG 675
QY 301 CARGATGMSNAATNACNGTNTYTAAYGTATACNTYTAAYGAYNSGANTNTAYACNTGY 360
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 676 CAGGAGGTGTCCATCTGCTCAACGTCACCTTAACACACTGACCTGACCTGACCTG 735
QY 361 AAYGTMSNMNGARTTYGARTTYGARGCNCAYMNCNTTYGTNAARAACNACMNGNTN 420
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 736 AATGTCTCCCGGAGTTTGATTTGAGGCGCATCGGCCCTTTGTGAAGACGACGGCTG 795
QY 421 ATNCCNYTMNGTNAACNGARGGCGNGARGATYTAACMSNGTNGTMSNGARATN 480
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 796 ATCCCCCTAAGAGTCAACGAGGAGGCTGAGAGACTTCACTCTGTGTCTCAGAAATC 855
QY 481 ATGATGTATATNTYNTNGTNTTYTNAACNTYNTGYTNTATNGARATNTATYTG 540
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 856 ATGATGTATACCTCTGTGTCTCTCCTCACCTGTGTGTCTCATGATGATATATTTGC 915
QY 541 TAYMNAARGTMSNAARGCNGARGGCGNGARGARAYGAGMSNGAYTAYTNGCN 600
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 916 TACAGAAAGTCTCAAAAGCCGAGAGGCGACGCCAAGAAACCGGTCTGACTACCTTGC 975
QY 601 ATNCCMSNGARAAAYARGARAAAYMSNGCNTNCCNGTNGARGAR 645
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 976 ATCCCATCTGAGAACAGAGAACTCTGCGGTACCACTGAGGAA 1020

```

RESULT 2

```

US-10-029-191-21
; Sequence 21, Application us/10029191
; Publication No. US2002016045A1
; GENERAL INFORMATION:
; APPLICANT: CURTIS, ROY A. J.
; TITLE OF INVENTION: NOVEL GENE ENCODING A SODIUM CHANNEL BETA-3 SUBUNIT
; FILE REFERENCE: 210147.00XX/5U1
; CURRENT FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 09/569,978
; PRIOR FILING DATE: 2000-05-12
; PRIOR APPLICATION NUMBER: US 60/134,198
; PRIOR FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 21
; LENGTH: 645
; TYPE: DNA
; ORGANISM: Rattus sp.
US-10-029-191-21

```

```

Query Match 71.8%; Score 462.8; DB 13; Length 645;
Best Local Similarity 58.4%; Pred. No. 4.9e-120; Indels 0; Gaps 0;
Matches 377; Conservative 145; Mismatches 123;

```

```

QY 1 ATGCCGACNTTAAVWNGTNTTYCCNYTNGCNSNYTNGTNTATNTAYTGGTMSN 60
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1 ATGCCGACCTTCAACAGATTTGCTTCCCTTACTTCTATGTCTCATCTACGTGGTCA 60
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 61 GNTGTCTTCGCTGTGTGTGAGAGTNGCNSNGARACNGARGCNGTNGARGNAAYCCN 120
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 61 GTCGTCTTCCTGTGTGTGTGAGAGTNGCCTCGAGACAGAAACGGGTGACGGGCAATCCC 120
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

```

QY 121 ATGAATYTMNGTGYATNWSNTGYATGAARMNGARGGTNGARGCNAACNCTNGTN 180
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 121 ATGAACCTCGCGCTCATCTCTGCATGAAGAGAGAGAGGTGAGGCCACACGTGTGGT 180
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 181 GARTGTTTAAVWNCNGARGGNGNAAAGATTTTNTATNTAYGARTAYMGNAAYGN 240
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 181 GATGTTTCTACAGGCGCTTACAGGCGGTAAAGATTTCTTATATATGATATCGAATGGC 240
QY 241 CAYCARGGTNGARGNSCCTTYTCARGGNGNTYTCARTGAAYGAGNSNAARGAYTN 300
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 241 CACCAAGAGGTGAGAGCCCTTTCAGGGCGCCTCAGATGGAATGACAGCAAGACCTG 300
QY 301 CARGATGMSNAATNACNGTNTYTAAYGTATACNTYTAAYGAYNSGANTNTAYACNTGY 360
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 301 CAGGAGGTGTCCATCTGCTCAACGTCACCTTAACACACTGACCTGACCTGACCTG 360
QY 361 AAYGTMSNMNGARTTYGARTTYGARGCNCAYMNCNTTYGTNAARAACNACMNGNTN 420
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 361 AATGTCTCCCGGAGTTTGATTTGAGGCGCATCGGCCCTTTGTGAAGACGACGGCTG 420
QY 421 ATNCCNYTMNGTNAACNGARGGCGNGARGATYTAACMSNGTNGTMSNGARATN 480
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 421 ATACCTTTCGAGTCACTGAAAGGCGGAGAAAGACTTCACTCTGTGTCTCGAAATC 480
QY 481 ATGATGTATATNTYNTNGTNTTYTNAACNTYNTGYTNTATNGARATNTATYTG 540
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 481 ATGATGTATACCTCTGTGTCTCTCCTCACCTGTGTGTGTATTAATGATGATCTATTTGC 540
QY 541 TAYMNAARGTMSNAARGCNGARGGCGNGARGARAYGAGMSNGAYTAYTNGCN 600
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 541 TACAGAAAGTCTCAAAAGCCGAGAGGCGACGCCAAGAAACCGGTCTGACTACCTTGC 600
QY 601 ATNCCMSNGARAAAYARGARAAAYMSNGCNTNCCNGTNGARGAR 645
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 601 ATCCCTTCAAGAGAACAGAGAACTCTGTGTGATCACTGTGAGGAA 645

```

RESULT 3

```

US-09-977-579-3
; Sequence 3, Application US/09977579
; Publication No. US20040248240A1
; GENERAL INFORMATION:
; APPLICANT: Cambridge University Technical Services
; TITLE OF INVENTION: A novel family of beta sub-unit proteins from a voltage gated sodi
; FILE REFERENCE: 674558-2001
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: PCT/EP00/01783
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/129,473
; PRIOR FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 3
; LENGTH: 2220
; TYPE: DNA
; ORGANISM: rat
US-09-977-579-3

```

```

Query Match 71.8%; Score 462.8; DB 11; Length 2220;
Best Local Similarity 58.4%; Pred. No. 1.4e-119; Indels 0; Gaps 0;
Matches 377; Conservative 145; Mismatches 123;

```

```

QY 1 ATGCCGACNTTAAVWNGTNTTYCCNYTNGCNSNYTNGTNTATNTAYTGGTMSN 60
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 363 ATGCCGACCTTCAACAGATTTGCTTCCCTTACTTCTATGTCTCATCTACGTGGTCA 422
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 61 GNTGTCTTCGCTGTGTGTGAGAGTNGCNSNGARACNGARGCNGTNGARGNAAYCCN 120
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 423 GTCGTCTTCCTGTGTGTGTGAGAGTNGCCTCGAGACAGAAACGGGTGACGGGCAATCCC 482
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

| | | | |
|----|-----|---|------|
| Qy | 12 | ATGAATYTMNGTATYATMWSNTGYTGAAPMGNBARGTNGARCNACNACGTGTGN | 1 80 |
| Db | 198 | ATGAACTAGGTGCATCTCTCGCATGAAGAGGAGAGGTGAGGCCACCATCTGTGTG | 2 57 |
| Qy | 181 | GAATGTTTATYAMGCCNGARGGNGNARGAATTTTATNTATYAGATYMGNAAYGN | 2 40 |
| Db | 258 | GAATGTTTCTACAGGCTCGAGGGCCGTAAATATTCCTTATATATYAGTATCGAAATGGC | 3 17 |
| Qy | 241 | CATCARGATNGARMSNCNTTTCARAGANNNTNCAFTGGAAYGGMSNAARGAYTN | 3 00 |
| Db | 318 | CACCAAGAAAGTGAGAGGCCCTTTCCAAAGCCGTCTGCATGTGAAATGGAGCAAAAGCTG | 3 77 |
| Qy | 301 | CARAGYTNMSNATNACNGTNTYNAAYGTNACNTNAAAYGAYWSNGAYTNAYACNTGY | 3 60 |
| Db | 378 | CAGAGGTATCATCACTGATCTCAATGTCACTTTGATGATCACTGTGGCCCTCTACACATGC | 4 37 |
| Qy | 361 | AAATGYSMNGARFTTGTGATTTGATTCARBCNCAVMGNCNTTGTNBARACNACMGAYTN | 4 20 |
| Db | 438 | AAATGTTCAGAGAGTTCGAATTCAGAGCAACAGGCCCTTTTGTGAAGACCAAGAGCTG | 4 97 |
| Qy | 421 | ATNCCNYTNMGNTNACNGARGARGCNGNGARGAATTTYACMSNGTNGTWSNGARATN | 4 80 |
| Db | 498 | ATACCTTTCCGAGTCACTGPAAGAGCGGGAAGACCTTCACTCCGTGCTCGGAATC | 5 57 |
| Qy | 481 | ATGATGTAAATYNTYNTGTNTTTTNTACNTNTNGYNTNTNATNGBARATGATNTATGY | 5 40 |
| Db | 558 | ATATGTATCAATCCTCTGCTCTTCTCTCACTTGTGGCTGTTTATTTGAAATGATCTATTGC | 6 17 |
| Qy | 541 | TATMGNAARGTWSNNAARGCNGARGCNGCNGARGAAPAYGCMWSNGAYTAYTNGN | 6 00 |
| Db | 618 | TACAGAAAGTCTCTAAGGCCGAAAGGCAACACAGGAAATGCGTCTGCATCACTTGCT | 6 77 |
| Qy | 601 | ATNCCMSNGARAAAYARGAAYMSNCCNCTNCCNCTNGARGAR | 6 45 |
| Db | 678 | ATCCCTTCAGAAACAGAGAACTCTGTGTACTCTGTGAGGGA | 7 22 |

```

RESULT 5
US-10-029-191-1
Sequence 1, Application US/10029191
Publication No. US20020160453A1
GENERAL INFORMATION:
APPLICANT: CURTIS, ROY A. J.
TITLE OF INVENTION: NOVEL GENE ENCODING A SODIUM CHANNEL BETA-3 SUBUNIT
TITLE OF INVENTION: PROTEIN
FILE REFERENCE: 210147.00XX/5U1
CURRENT APPLICATION NUMBER: US/10/029,191
CURRENT FILING DATE: 2001-12-20
PRIOR APPLICATION NUMBER: 09/569,978
PRIOR FILING DATE: 2000-05-12
PRIOR APPLICATION NUMBER: US 60/134,198
PRIOR FILING DATE: 1999-05-14
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 3108
TYPE: DNA
ORGANISM: Rattus sp.
US-10-029-191-1

```

Query Match 71.8%; Score 462.8; DB 13; Length 3108;
Best Local Similarity 58.4%; Pred. No. 1.8e-119;
Matches 377; Conservative 145; Mismatches 123; Indels 0; Gaps 0;

```

Db 198 ATGAGCTGAGTGCATCTCTCATGAGGAGAGAGTGGAGCCCACTGTGGTG 257
Qy 181 GARTGTTTATYMGNCNGARGNGNAAAGATTTTNAATNTAYGARTYMGNAAYGN 240
Db 258 GAGTGTCTTACAGGCGCTGAGGGCGGTAAAGATTTCTATATATATAGATCGAAATGCG 317
Qy 241 CAYCARGARTNGARBSCNTTTCARGGMMGYTCARTGAAAYGAGNMSNARGAYTN 300
Db 318 CACGAGAAAGTGAAGAGCCCTTCCAGAGCCGTCTGACGTGGAATGGAGCAAAAGCTCG 377
Qy 301 CARGAYTNMSNATNACNGTNTYNAAYGTNACNTYNAAYGAYVNSNGYNTNTAYACNTGY 360
Db 378 CAGGAGTATCCATCACTGATCAATGTCACATTGAATGACTCTGGCCCTCAACATGCG 437
Qy 361 AAYGTMSMNGARNTTYGARTTYGARGNCAMGNCNTTYTNAARAACNMGNTYN 420
Db 438 AATGTGTCCAGGAGGTTCGAAATTCGAGGACACAGGCTTTTGTGAAGCAACGAGACTG 497
Qy 421 ATNCCNYTMNGTNCNGARGARGNGNGARGATTTTAAACNMSNGTNGTMSNGARATN 480
Db 498 ATACCTTTGGGAGTCACTGAAGAGCGGAGAGAACTTCACTCCGTGCTTGGAAATC 557
Qy 481 ATGATGTAAYATYNTYNTGNTTNTYTNACNTYNTGYTYNTATNGARATGATNTATGY 540
Db 558 ATGATGATACATCTCTGCTGCTTCTCTCACTGTGCTGTTATGAGATGATCTATTCG 617
Qy 541 TAYMGAAAGTMSNAAARGCNGARGARGCNGCNCARGAARAYCNCMSNATYTYTNGCN 600
Db 618 TACAGAAAGTCTCTAAGGCCGGAAGAGGACGACACAGAAATATGCGTCACTACCTTGCT 677
Qy 601 ATNCCMSNGARAAAYARGARAAAYMSNGCNGTNCNGTNGARGAR 645
Db 678 ATCCCTTCAGAGAACAGAGAACTCTGTGTGTAACCTGTGAGAGAA 722

```

RESULT 6

```

US-10-723-860-2247
; Sequence 2247, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aaliz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnick, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; PRIOR FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2247
; LENGTH: 1335
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-723-860-2247

```

Query Match 24.6%; Score 158.4; DB 18; Length 1335;
 Best Local Similarity 38.2%; Pred. No. 1.5e-33;
 Matches 240; Conservative 101; Mismatches 266; Indels 21; Gaps 3;

```

Qy 25 CCNYTNGCWSNYTNGTNTYNTATNTAYTGGTMSNGTNTGYTTCNCGNTNGYTNARG 84
Db 34 CTGCTGGCCCTTATGTTGCGCGCGGCACTGGTGTCTCAGCCTGCGGGGCGTGGAG 93
Qy 85 GTNCCMSNGARACNARGCNGTNCARGNAAYCCNATGARATYTNMGNTGYATNMSNTGY 144
Db 94 GTGGAATCGAAGCCGAGGCCGTGTATGGATACCTTCAAAATTTCTTGGCATCTCTGCG 153
Qy 145 ATGAARMGNGARGARTNGARGCNCACNGTNGTNGARTGTTTAYMGNCNGARGN 204

```

```

Db 154 AAGCGCCGAGGAGACCAACCTGAGACCTTCAACGAGTGAACCTTCCGCCAAGAGGCG 213
Qy 205 GGNAAAGATTTTNTA-----TNTAYGARTYMGNAAYGAGNAYCARGARGTNGAR 255
Db 214 ACTGAGGAGTTTGTCAAGATCTCGCGCTAYGAGAAATGAGGTGTTGACGTGAGAGGAGAT 273
Qy 256 MSNCNTTTCARGMMNGYNTNCARTGAAY-----GMSNNAARGAAYTNCARGAY 306
Db 274 GAGCGCTTGAAGGGCGCGGTGTGTGAATGCAAGCGGGGACCAAAACCTCGACGAGAT 333
Qy 307 GTNMSNATNACNGTNTYNAAYGTNACNTYNAAYGAYVNSNGYNTNTAYACNTGYATN 366
Db 334 CTGTCTATCTTCATCAACCATGTCACTCAACCACTCGGGGACATCAACGAGCCACGTC 393
Qy 367 MSNMGARTTYGARTTYGARGONCAVMGNCNTTGTNABACNACNMGNTNATNCCN 426
Db 394 TACCGCTCTCTTCTTTCGAAAACCTACGAGCACACACACAGCGTCTCAAGAAATCCAC 453
Qy 427 YTNMGNTNACNGARGARGCNGNGARGATTTTAAACNMSNGTNGTMSNARGATNATGATG 486
Db 454 ATTGAGTGTGAGCAAAACCAAGAGACATGCAATCATGTGTCTGATCATGATG 513
Qy 487 TAYATNTYNTGNTTNTYTNACNTYNTGYTYNTATNGARATGATNTATYTYTAYMGN 546
Db 514 TATGTGCTCATTTGTGTGTGAACATATGCTGTGCGAGATGATTTTACTGTCTAACG 573
Qy 547 AARGTMSNAAARGCNGARGAR---GNGCNCARGAARAAAYCNCMSNATYTYTNGCATN 603
Db 574 AAGATGCTGCGGCCACGAGACTGTGTCACAGAGAAATGCTCGGAATACCTGCGCATC 633
Qy 604 CCNMSNGARAAAYARGARAAAYMSNGCNG 631
Db 634 ACCTGTGAAGCAAGAGAACTGCAACGG 661

```

RESULT 7

```

US-10-029-191-23
; Sequence 23, Application US/10029191
; Publication No. US20020160453A1
; GENERAL INFORMATION:
; APPLICANT: CURTIS, ROY A.J.
; TITLE OF INVENTION: NOVEL GENE ENCODING A SODIUM CHANNEL BETA-3 SUBUNIT
; FILE REFERENCE: 210147.00XX/501
; CURRENT APPLICATION NUMBER: US/10/029,191
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 09/569,978
; PRIOR FILING DATE: 2000-05-12
; PRIOR APPLICATION NUMBER: US 60/134,198
; PRIOR FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 657
; TYPE: DNA
; ORGANISM: Lepus Sp.
US-10-029-191-23

```

Query Match 24.5%; Score 158; DB 13; Length 657;
 Best Local Similarity 38.1%; Pred. No. 1.1e-33;
 Matches 239; Conservative 102; Mismatches 266; Indels 21; Gaps 3;

```

Qy 25 CCNYTNGCWSNYTNGTNTYNTATNTAYTGGTMSNGTNTGYTTCNCGNTNGYTNARG 84
Db 10 CTGCTGGCTTTGTGTGCGCGCGCGGCTGTGTCTCGGCTGGGGGGGCTGCTGAG 69
Qy 85 GTNCCMSNGARACNARGCNGTNCARGNAAYCCNATGARATYTNMGNTGYATNMSNTGY 144
Db 70 GTGGAATCGAAGCCGAGAACCGGTGTACGGGATGACCTTCAAAATTTCTGTCATCTCTGCG 129
Qy 145 ATGAARMGNGARGARTNGARGCNCACNGTNGTNGARTGTTTAYMGNCNGARGN 204
Db 130 AAGCGCCGAGGAGACCAAGGCCGAGACCTTCAAGAGTGAACCTTCCGCCAAGAGGCG 189

```


| | | | | |
|------------------------|------------------|---|----------------|-------------|
| Query Match Similarity | 24.3% | Score 156.8 | DB 18 | Length 1414 |
| Best Local | Similarity 38.1% | Pred. No. 4.5e-33 | | |
| Matches | 239 | Conservative 101 | Mismatches 267 | Indels 21 |
| | | | Gaps | 3 |
| Qy | 25 | CCNNTGCMNSYNTNGTNTYATNTATYATGGGTMSNGTNTGYTTYCCNGTNTGVTNGAR | 84 | |
| Db | 110 | CTGCTGGCGCTTAGTGATGTCGGCGCGACATCGTGCTCTAGCCCTCGGGGGGCTGCTGGAG | 169 | |
| Qy | 85 | GTNCCMSNGSRAFCNGARGCNGTNGARCGMNAVCCNATGAARVTMMGTNGVATMNSNTGY | 144 | |
| Db | 170 | GTGAGCTCGGAGACCGAGGCGCGTGTATGGGATGAGACCTTCAMAAATCTTTGCAATCTTCCTGC | 229 | |
| Qy | 145 | ATGAARMGNGARGARGTNGARGCNA.CNACGTGTNGARTGTTYATYAMNCNGARGSN | 204 | |
| Db | 230 | AAGCCCCCAGCGAGAACCAACGCTAGACCTTCACCGAGTGACCTTCGCGCAGAAAGGC | 289 | |
| Qy | 205 | GGNARAGAYTTYTYNA-----TNTAYGARTYMYMNAAYGNCAYCARGARGSTNGAR | 255 | |
| Db | 230 | ACTGAGGAGTTTGTCAAGATCCTCGGCGCATNGAATGAGAGTGTTCGACCTGGAGGAGAT | 349 | |
| Qy | 256 | WSNCCNTTYCARGMNGVNTNCARTGAAV-----GGNWSNABAGAYTTCARAY | 306 | |

```

RESULT 9
US-09-917-800A-1654
; Sequence 1654, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
; APPLICANT: Mendick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Casle, Arthur
; APPLICANT: Elashoff, Michael
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5038-US
; CURRENT APPLICATION NUMBER: US/09/917,800A
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,645
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/292,336
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1654
; LENGTH: 1490
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURES:
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM_017288
US-09-917-800A-1654

Query Match      23.6%;   Score 152.2;   DB 9;   Length 1490;
Best Local Similarity  38.7%;   Pred No. 9; 5e-32;

```

Matches 232; Conservative 96; Mismatches 250; Indels 21; Gaps 3;

QY 54 GGTMSNGNTNTGTYTTCNGTNTGTNGARNTCCMSNGARACNGAGCNGTNCARB 113
DB 258 GGTATCTCCAGCTGGGGGGCTGCGTGGAGGTGTAATCTGAGACCGAGCGAGTATG 317
QY 114 NAAVCNATGAARYTMMGTGATYATMSNTGYATGARMGNGAGGTGARGCNCAN 173
DB 318 GATGACCTTCAAAATCTGTGTATCTCTGTAAGCCTCGTAGGACACCGCCGAGAC 377
QY 174 NGTNGTNGARTGTGTYTAYMNCNGARGNGNNAAGAVTYTNA-----TNTA 224
DB 378 CTTCACGAGGTGACCTTCGCGCAGAGGACACAGAGAAATTTGTCAAGTCTTAAGCTA 437
QY 225 YGARTYMGNAAYGNCAYCARGBTNGARMSCNTTYCARGMMGYTNCARTGAA 284
DB 438 TGAATAATGAGTGTCTCAGCTGAGAGAAATGAGCCCTTGGAGGCCGTGTGTGGA 497
QY 285 Y-----GGMNSNARGAAYTTCARGAYGTMSNATNACNGTYTNAAYGTNACNT 335
DB 498 CGGTACTCGGGGACCAAGAGCCTGACAGACCTGTCTCATCTTTCATCACCAGTGTCACT 557
QY 336 NAAVGA YMSNGNYTNTAYACNTGYA YGTMSNMNGARTTYGARTTYGARGCNCAYNG 395
DB 558 CAACCACTCTGGCGACTACGAATGTCA CGTCTACCGCTCTCTTGATTAATTAACA 617
QY 396 NCNTTYGTNAARACNACMNGNTNATNCNTNMNGTNGACGARGCNGNBARA 455
DB 618 GCACAAACCCAGCGCTGTCAAGAAATCCACTGAGAGGTGGTGAACAGGCCAACAGAG 677
QY 456 YTTTACNMSNGTNGTMSNGARATNATGATGATYATNTYNTGTYTNAACNTNTG 515
DB 678 TATGATCATCATGTGTGACAGATCATGATGATGATGATGATGATGATGATGATG 737
QY 516 GTTNTNATNGARTATGATYATGATYATGNAARGTMSNARF---CNGARGCNGC 572
DB 738 GCTCGTGGCGAGATGTGTACTGTACAGAAAGATGCTGCTGCCACCGAAGCTCTGC 797
QY 573 NCARGAARAYGCMWSNGAYATAYTNGCNAATCCNMSNGARAAAYARGARAYMSNCG 631
DB 798 ACAAGAAATGCTCGAATATCTGCGCATTACTTCGAGAGCAAGAGAACTGTACAG 856

RESULT 10
US-10-276-774-718/c
; Sequence 718, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y, Tom et al
; TITLE OF INVENTION: No. US20040053245A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: US/10/276,774
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 718
; LENGTH: 407
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-276-774-718

Query Match 16.3%; Score 105; DB 17; Length 407;
Best Local Similarity 58.0%; Pred. No. 7.4e-19;
Matches 87; Conservative 33; Mismatches 30; Indels 0; Gaps 0;

QY 444 RCGNGNGARGATYTTACMSNGTNGTMSNGARATNATGATYATNTYNTGNTT 503
DB 231 AGCTGAGAGAGACTTACCTGTGTGTCTCAGAAATCATGATGTATCATCTTCTGTGTC 172

QY 504 YTTNACNTNTGTYTNTYNTATNGARTGATNTATYTGTYTAYMNGARGTMSNARGCNGA 563
DB 171 CCTACCTTGCTGCTGCTCATGAGATGATATTTCTACAGAAAGTCTCAAAAAGCCCA 112
QY 564 RGARGCNGCARGARAAAYGCMWSNGAYTA 593
DB 111 AGAGGAGGCCCAAGAAACGCGTAAGTCCA 82

RESULT 11
US-10-401-916-12
; Sequence 12, Application US/10401916
; Publication No. US20040002439A1
; GENERAL INFORMATION:
; APPLICANT: Qin, Ning
; APPLICANT: Codd, Ellen
; APPLICANT: D'Andrea, Michael
; TITLE OF INVENTION: DNAS encoding human beta1a sodium channel subunit
; FILE REFERENCE: ORT-1221
; CURRENT APPLICATION NUMBER: US/10/401,916
; PRIOR FILING DATE: 2003-03-28
; PRIOR APPLICATION NUMBER: US/09/875,456A
; PRIOR FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 807
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-401-916-12

Query Match 12.6%; Score 81.2; DB 17; Length 807;
Best Local Similarity 33.5%; Pred. No. 7.3e-12;
Matches 157; Conservative 72; Mismatches 222; Indels 18; Gaps 2;

QY 25 CCNTYNGCMWSNYTNGTNTYNTATNTATYTGGTMSNGTNTGTYTTCNGTNTGTNGAR 84
DB 10 CTGCTGCGCTTATGTGTGCGCGGCGACATGGTGTCTCAGCCTGCGGGGCTGCGTGAG 69
QY 85 GTNCCMSNGARACNARGCNGTNCARGNAAYCCNATGAARTYNTMGTGYATMSNTGY 144
DB 70 GTGACTCGAGACCGAGGCCGTGTATGGATGACCTTCAAAATCTTTGCAATCTCTGC 129
QY 145 ATGAARMNGARGARTNGARGTNARGCNACNACNGTNGTNGARTGTTTYAYMNGCNGARG 204
DB 130 AAGCGCCGACGAGACCAACGCTGAGACCTTCACCGAGTGAACCTTCGCCGCAAGAGGC 189
QY 205 GGNARGAAYTYTNA-----TNTAYGARTAYMNAAYGNCAYCARGARTNGAR 255
DB 190 ACTGAGAGATTGTCAAGATCTCTGCGCTATGAGAAATGAGGTGTTGACGCTGAGAGGAT 249
QY 256 WSNCCCTTYCARGMNGNTNARGGAAY-----GGMNSNARGAAYTNCARGAY 306
DB 250 GAGCGCTTGAGGGCGCGGTGTGTGAATGCGAGCGGGGACCAAGACCTGCAAGAT 309
QY 307 GTMSNATNACNGTNTYNTAAYGTNACNTYNTAAYGAYMSNGNYTNTAYACNTGYAAYTN 366
DB 310 CTGTCTATCTTCATACCAATGTCACTCAACCACTCGGCGACATCAGAGTCCACGTC 369
QY 367 WSNMNGARTTYGARTTYGARGCNCAYMNCNTTYGTNAABACNACMNGNTNATNCCN 426
DB 370 TACCGCTGCTCTCTTTCGAAAACCTACGAGCACACACAGCGCTCAAGAAATCCAC 429
QY 427 YTMNGTNAACNARGARCGCNGNARGAYTYTACMSNGTNGTMSNG 475
DB 430 ATTGAGTGTGACAAAGTAGTGTGCGGTCTGCTGCCCTTTACCG 478

RESULT 12
US-10-401-916-13
; Sequence 13, Application US/10401916
; Publication No. US20040002439A1


```
QY 276 NCARTGGAAYGGMWSNAARGAAYTTCNARGAYGTNMSNATNACNGTNTYNAAYGTCNACNT 335
Db 3442 TGGGGGTGATTTTATTAGTTTATTAGTTGGGATTAATTAATTAATTAATTAATTAATTAAT 3501
QY 336 NAAVGAYWSNGNYTNTAYACNTGYAAYGTNMSNMNGARTTYGARTTYGARGCNCAYMG 395
Db 3502 TAGGTTAGGAGGTTGAGTTTATTATTAAGTTTATATAGTTAAGTTTATAGTTATATTA 3561
QY 396 NCCNTTYGNNARACNACNMGNNTAATNCCNYTNMGNGTNCNARGANGCNGANGARGA 455
Db 3562 GGTGTTAGTGTGTTGGGTTGTTTATTTTATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTG 3621
QY 456 YTTTACNMSNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNG 515
Db 3622 TATGATTTATTTGTTGTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 3681
QY 516 GYTNNTATNGARATGATNTATYGTAYTMGNNAARGT 551
Db 3682 GTTCGTGTAGAGATGATTTATTTATTAAGAAGAT 3717
```

```
RESULT 15
US-10-184-644-346
; Sequence 346, Application US/10184644
; Publication No. US20030044930A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Goddard, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C227
; CURRENT APPLICATION NUMBER: US/10/184,644
; PRIOR APPLICATION DATE: 2002-06-28
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 346
; LENGTH: 671
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-184-644-346
```

Query Match 7.8%; Score 50; DB 14; Length 671;
Best Local Similarity 5.0%; Pred. No. 0.0043;
Matches 31; Conservative 193; Mismatches 394; Indels 0; Gaps 0;

```
QY 28 YTNCGMWSNYTNGTNTATNTATYATGCGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNG 87
Db 11 VPAKMGYHPHWPARIIDIDGAVKPPNKKYPIFFPGTHETAFILGPKDLPPYDKCKDXGK 70
QY 88 CCNWSNGARACNARGCNGTNCARGGNAAYCNATGATNTNGTNGTNGTNGTNGTNGTNGTNG 147
Db 71 PNRKGFNEGLMEIQNNPHASYSAPPPVSSDSEAPANPADSDADEDEDEDEGVAVTA 130
QY 148 AARMGNGARGARGTNGARGCNAACNCGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNG 207
Db 131 VTATTAASDRHESDSDSDSGNLKRTPALMSVSKRARKASDDQASVSPSEENS 190
QY 208 AARGATTTTATNTATNTAYGARTAYMGAAAYGNCAYCARGARGTNGARWSNCCNTTYCAR 267
Db 191 ESSSESEKTSDDQFTPEKKAAYAPRPLGRKKKKKAPASDSKADSDAKPEPVAM 250
QY 268 GGMGNYTNCARTGGAAYGGMWSNAARGAAYTTCNARGAYGTNMSNATNACNGTNTYNAAY 327
```

```
Db 251 ARSASSSSSSSDSDSVKPPRGRKPAEKPLPKPRGRKPKPERPPSSSSSDSDSDDEV 310
QY 328 GTNACNYTAAAGAYWSNGNTNTATYACNTGYAAYGTNWSNMNGARTTYGARTTYGAR 387
Db 311 DRISWKRRDEARRRRELEARRRREDEELRLRLEOEKEKEKERRERREADGAEARGSGSS 370
QY 388 GCNCAYMNCNTTYGTNARACNACNMGNNTAATNCCNYTNMGNGTNCNARGARGCN 447
Db 371 GDELRDEDPVKKRGRKGRGPSSSDSEPEALEREAKSKAKKQSSSTBPARKPGOK 430
QY 448 GNGARGAAYTTCNWSNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNG 507
Db 431 EKRVRPEERQOAKPVYERTRRSEBFSMDRVEKEKEPSVEBKQKLSHSEIKFALKVDS 490
QY 508 ACNYTNGTNTYNTATNGARATGATNTATYGTAYTMGAARGTNGWSNARCGNGARGAR 567
Db 491 PDVKRCLNLEBELGTLQVTSQILQKNTDVATLTKIRRYKANADVKEKAAEYVTRLKSHV 550
QY 568 GCNCGNCARGAARAAAYGCMWSNGAYTAYTNGCNATNCCNWSNGARAAAYARGAARAAWSN 627
Db 551 LBPXLEAYGVKVKAGMEKEKEKAEKLEAGBELAGBEAPQEKAEKPSYDLSAPVNGEATSQK 610
QY 628 GCNCGTNCNGTNGARGAR 645
Db 611 GBSAEDKEHEBGRDSEEG 628
```

Search completed: April 7, 2005, 23:40:37
Job time : 504 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2005, 21:17:03 ; Search time 3108 Seconds

(without alignments)

7899,438 Million cell updates/sec

Title: US-09-977-579-2-REV

Sequence: 1 ATGCCGNCNTTAAAYMGNNT.....SNGCNGTNCNGTNGARGAR 645

Scoring table: IDENTITY/NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST: *
1: gb_est1.*
2: gb_est2.*
3: gb_nuc.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gss1.*
9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|-----------|----------------------|
| 1 | 466 | 72.2 | 648 | 9 | AY419145 | AY419145 Homo sapi |
| 2 | 466 | 72.2 | 2555 | 3 | CR609664 | CR609664 full-length |
| 3 | 466 | 72.2 | 4052 | 3 | HSMB01563 | AL136589 Homo sapi |
| 4 | 465.6 | 72.2 | 1062 | 5 | BX420015 | BX420015 BX420015 |
| 5 | 461.2 | 71.5 | 648 | 9 | AY419147 | AY419147 Mus muscu |
| 6 | 461.2 | 71.5 | 1359 | 3 | AK076466 | AK076466 Mus muscu |
| 7 | 461.2 | 71.5 | 3549 | 3 | AK049747 | AK049747 Mus muscu |
| 8 | 450.2 | 69.8 | 4149 | 3 | AK049286 | AK049286 Mus muscu |
| 9 | 447.8 | 69.4 | 636 | 5 | BM933157 | BM933157 UI-M-BH3- |
| 10 | 447.8 | 69.4 | 975 | 5 | BX445002 | BX445002 BX445002 |
| 11 | 442 | 68.5 | 672 | 2 | BB614118 | BB614118 BB614118 |
| 12 | 433.4 | 67.2 | 4105 | 3 | BC058083 | BC058083 Mus muscu |
| 13 | 405.2 | 62.8 | 950 | 4 | BG294174 | BG294174 603391245 |
| 14 | 395 | 61.2 | 582 | 4 | BP200910 | BP200910 BP200910 |
| 15 | 385 | 59.7 | 723 | 6 | CA749311 | CA749311 UI-M-FY0- |
| 16 | 351.8 | 54.5 | 927 | 5 | BQ713131 | BQ713131 AGENCOURT |
| 17 | 348 | 54.0 | 986 | 5 | BUI18914 | BUI18914 603142291 |
| 18 | 342 | 53.0 | 584 | 7 | AY419146 | AY419146 Pan trogl |
| 19 | 341.4 | 52.9 | 682 | 7 | CN219870 | CN219870 WLA010B08 |
| 20 | 333.2 | 51.7 | 584 | 5 | BP361278 | BP361278 BP361278 |
| 21 | 328.4 | 50.9 | 825 | 5 | BQ745919 | BQ745919 UI-M-EXO- |
| 22 | 324.4 | 49.9 | 846 | 6 | CD355879 | CD355879 UI-M-FY0- |
| 23 | 321.6 | 49.9 | 892 | 5 | BUI36814 | BUI36814 603790206 |
| 24 | 320.6 | 49.7 | 595 | 5 | BUI368531 | BUI368531 603788820 |

| | | | | | | |
|----|-------|------|------|---|----------|---------------------|
| 25 | 314.2 | 48.7 | 895 | 5 | BX743834 | BX743834 BX743834 |
| 26 | 309.8 | 48.0 | 742 | 6 | CB526211 | CB526211 UI-M-FY0- |
| 27 | 307.6 | 47.7 | 582 | 5 | BP311484 | BP311484 BP311484 |
| 28 | 293 | 45.4 | 845 | 6 | CA327438 | CA327438 UI-M-FY0- |
| 29 | 291 | 45.1 | 798 | 7 | CF290312 | CF290312 AGENCOURT |
| 30 | 288 | 44.7 | 700 | 4 | B1739617 | B1739617 603361873 |
| 31 | 283 | 43.9 | 652 | 2 | BB652801 | BB652801 BB652801 |
| 32 | 275.4 | 42.7 | 1069 | 6 | CB787935 | CB787935 AMGNNTUC:N |
| 33 | 266.2 | 41.3 | 823 | 7 | BM928131 | BM928131 AGENCOURT |
| 34 | 263 | 40.8 | 734 | 7 | CO428886 | CO428886 UI-M-HXO- |
| 35 | 254.8 | 39.5 | 734 | 7 | CK367344 | CK367344 AGENCOURT |
| 36 | 254.2 | 39.4 | 712 | 6 | CB526257 | CB526257 UI-M-FY0- |
| 37 | 248.4 | 38.5 | 780 | 5 | BQ770528 | BQ770528 UI-M-FY0- |
| 38 | 243 | 37.7 | 750 | 6 | CD349206 | CD349206 UI-M-FY0- |
| 39 | 235.2 | 36.5 | 714 | 7 | CF531573 | CF531573 UI-M-FY0- |
| 40 | 233.8 | 36.2 | 901 | 7 | CF591773 | CF591773 AGENCOURT |
| 41 | 232.6 | 36.1 | 669 | 5 | BM951151 | BM951151 UI-M-EHO- |
| 42 | 221.6 | 34.4 | 471 | 6 | CB732717 | CB732717 AMGNNTUC:N |
| 43 | 218 | 33.8 | 544 | 7 | CN666608 | CN666608 A0842P07- |
| 44 | 218 | 33.8 | 650 | 7 | CN665870 | CN665870 A0832C11- |
| 45 | 217.4 | 33.7 | 582 | 5 | BP202832 | BP202832 BP202832 |

ALIGNMENTS

| | | | | | |
|------------|--|--------|-----|--------|-----------------|
| RESULT 1 | AY419145 | 648 bp | DNA | linear | GSS 12-DEC-2003 |
| LOCUS | AY419145 | | | | |
| DEFINITION | AY419145 | | | | |
| ACCESSION | AY419145 | | | | |
| VERSION | AY419145.1 | | | | |
| KEYWORDS | GSS. | | | | |
| SOURCE | Homo sapiens (human) | | | | |
| ORGANISM | Homo sapiens | | | | |
| REFERENCE | Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B., Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. | | | | |
| AUTHORS | Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B., Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. | | | | |
| TITLE | Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA | | | | |
| JOURNAL | Science 302 (5652), 1960-1963 (2003) | | | | |
| PUBMED | 14671302 | | | | |
| REFERENCE | 2 (bases 1 to 648) | | | | |
| AUTHORS | Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B., Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. | | | | |
| TITLE | Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA | | | | |
| JOURNAL | Science 302 (5652), 1960-1963 (2003) | | | | |
| PUBMED | 14671302 | | | | |
| REFERENCE | 2 (bases 1 to 648) | | | | |
| AUTHORS | Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B., Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. | | | | |
| TITLE | Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA | | | | |
| JOURNAL | Science 302 (5652), 1960-1963 (2003) | | | | |
| PUBMED | 14671302 | | | | |
| REFERENCE | 2 (bases 1 to 648) | | | | |
| AUTHORS | Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B., Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. | | | | |
| TITLE | Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA | | | | |
| JOURNAL | Science 302 (5652), 1960-1963 (2003) | | | | |
| PUBMED | 14671302 | | | | |
| REFERENCE | 2 (bases 1 to 648) | | | | |
| AUTHORS | Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B., Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. | | | | |
| TITLE | Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA | | | | |
| JOURNAL | Science 302 (5652), 1960-1963 (2003) | | | | |
| PUBMED | 14671302 | | | | |
| REFERENCE | 2 (bases 1 to 648) | | | | |
| AUTHORS | Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B., Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. | | | | |
| TITLE | Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA | | | | |
| JOURNAL | Science 302 (5652), 1960-1963 (2003) | | | | |
| PUBMED | 14671302 | | | | |
| REFERENCE | 2 (bases 1 to 648) | | | | |
| AUTHORS | Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B., Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. | | | | |
| TITLE | Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA | | | | |
| JOURNAL | Science 302 (5652), 1960-1963 (2003) | | | | |
| PUBMED | 14671302 | | | | |
| REFERENCE | 2 (bases 1 to 648) | | | | |
| AUTHORS | Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B., Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. | | | | |
| TITLE | Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA | | | | |
| JOURNAL | Science 302 (5652), 1960-1963 (2003) | | | | |
| PUBMED | 14671302 | | | | |
| REFERENCE | 2 (bases 1 to 648) | | | | |
| AUTHORS | Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B., Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. | | | | |
| TITLE | Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA | | | | |
| JOURNAL | Science 302 (5652), 1960-1963 (2003) | | | | |
| PUBMED | 14671302 | | | | |
| REFERENCE | 2 (bases 1 to 648) | | | | |
| AUTHORS | Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B., Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. | | | | |
| TITLE | Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA | | | | |
| JOURNAL | Science 302 (5652), 1960-1963 (2003) | | | | |
| PUBMED | 14671302 | | | | |
| REFERENCE | 2 (bases 1 to 648) | | | | |
| AUTHORS | Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B., Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. | | | | |
| TITLE | Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA | | | | |
| JOURNAL | Science 302 (5652), 1960-1963 (2003) | | | | |
| PUBMED | 14671302 | | | | |
| REFERENCE | 2 (bases 1 to 648) | | | | |
| AUTHORS | Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B., Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. | | | | |
| TITLE | Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA | | | | |
| JOURNAL | Science 302 (5652), 1960-1963 (2003) | | | | |
| PUBMED | 14671302 | | | | |
| REFERENCE | 2 (bases 1 to 648) | | | | |
| AUTHORS | Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B., Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. | | | | |
| TITLE | Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA | | | | |
| JOURNAL | Science 302 (5652), 1960-1963 (2003) | | | | |
| PUBMED | 14671302 | | | | |
| REFERENCE | 2 (bases 1 to 648) | | | | |
| AUTHORS | Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B., Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. | | | | |
| TITLE | Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA | | | | |
| JOURNAL | Science 302 (5652), 1960-1963 (2003) | | | | |
| PUBMED | 14671302 | | | | |
| REFERENCE | 2 (bases 1 to 648) | | | | |
| AUTHORS | Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B., Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. | | | | |
| TITLE | Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA | | | | |
| JOURNAL | Science 302 (5652), 1960-1963 (2003) | | | | |
| PUBMED | 14671302 | | | | |
| REFERENCE | 2 (bases 1 to 648) | | | | |
| AUTHORS | Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B., Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. | | | | |
| TITLE | Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA | | | | |
| JOURNAL | Science 302 (5652), 1960-1963 (2003) | | | | |
| PUBMED | 14671302 | | | | |
| REFERENCE | 2 (bases 1 to 648) | | | | |
| AUTHORS | Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B., Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. | | | | |
| TITLE | Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA | | | | |
| JOURNAL | Science 302 (5652), 1960-1963 (2003) | | | | |
| PUBMED | 14671302 | | | | |
| REFERENCE | 2 (bases 1 to 648) | | | | |
| AUTHORS | Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B., Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. | | | | |
| TITLE | Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA | | | | |
| JOURNAL | Science 302 (5652), 1960-1963 (2003) | | | | |
| PUBMED | 14671302 | | | | |
| REFERENCE | 2 (bases 1 to 648) | | | | |
| AUTHORS | Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B., Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. | | | | |
| TITLE | Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA | | | | |
| JOURNAL | Science 302 (5652), 1960-1963 (2003) | | | | |
| PUBMED | 14671302 | | | | |
| REFERENCE | 2 (bases 1 to 648) | | | | |
| AUTHORS | Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B., Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. | | | | |
| TITLE | Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA | | | | |
| JOURNAL | Science 302 (5652), 1960-1963 (2003) | | | | |
| PUBMED | 14671302 | | | | |
| REFERENCE | 2 (bases 1 to 648) | | | | |
| AUTHORS | Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B., Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. | | | | |
| TITLE | Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA | | | | |
| JOURNAL | Science 302 (5652), 1960-1963 (2003) | | | | |
| PUBMED | 14671302 | | | | |
| REFERENCE | 2 (bases 1 to 648) | | | | |
| AUTHORS | Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B., Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. | | | | |
| TITLE | Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA | | | | |
| JOURNAL | Science 302 (5652), 1960-1963 (2003) | | | | |
| PUBMED | 14671302 | | | | |
| REFERENCE | 2 (bases 1 to 648) | | | | |
| AUTHORS | Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B., Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. | | | | |
| TITLE | Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA | | | | |
| JOURNAL | Science 302 (5652), 1960-1963 (2003) | | | | |
| PUBMED | 14671302 | | | | |
| REFERENCE | 2 (bases 1 to 648) | | | | |
| AUTHORS | Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civeille,D.R., Lu,F., Murphy,B., Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M. | | | | |
| TITLE | Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA | | | | |
| JOURNAL | Science 302 (5652), 1960-1963 (2003) | | | | |
| PUBMED | 14671302 | | | | |

| | | | |
|----|-----|---|-----|
| QY | 6 | GTNNGTYYCCNGNNMTGTGNTGABRTNCCMNSNGBARCNGARGCNGTTCAGGNAAYCCN | 120 |
| Db | 61 | GTCTGCTCCCTGTGTGTGTGAATGCGCCCTGGAGAGAGGCGCTGTGAGGGCAACCCC | 120 |
| QY | 121 | ATGAARYTNMGNTGYATWMSNTGYATGAABMGNGARGRTNGARGCNAACNACNTNGTN | 180 |
| Db | 121 | ATGAAGCTGCGCTGCATCTCTGCATGAAGAAGAGAGGTGAGGCCACACGCTGTGTG | 180 |
| QY | 181 | GARTGGTYYTAYMGNCCNGARGNGNGNARGAYYTTYTNATNTAYGARTAYMGNNAAYGN | 240 |
| Db | 181 | GAATGTCTCTACAGGCCCGAGGGCGGTAAAGATYTCCTTAATTTMGAGATATCGAATGGC | 240 |
| QY | 241 | CAYCARBGRNNGARWMSNCNTTTCARBGNNNTNRCARTGGAAYGNNMNSNAARAYTN | 300 |
| Db | 241 | CACCAAGAGGTGAGAGGCCCTCTTCAGGGGGCGCTCAGTGGAAATGACAGCAAGGACCTG | 300 |
| QY | 301 | CARGAYGTNMSNATNACNGTYNTNAAAYGTNACNTYNAAYGAYWNSNGNYTNAYACNTGY | 360 |
| Db | 301 | CAGAACGTGTCAATCACTGTGCTCAAGCTCACTCGAAGACTCTGGCGCTCTACACCTGC | 360 |
| QY | 361 | AAVGTNMSNMNGARTTYGARTTYGARGCNCAYMGNCCNTTYGTNAPACNAACMNGNTTN | 420 |
| Db | 361 | AAATGTGTCCCGGAGGTTTGAGTTTGAGGCGCATCGGCCCTTGTGTAAAGACGACGGCGTG | 420 |
| QY | 421 | ATTNCCNTNMNGNACNGARGARCGNNGNARGAYTYTAYCMWSNGTNGTMSNABARTN | 480 |
| Db | 421 | ATCCCCCTMAAGTCAACGAGGAGGCTGAGAGGACTTCACTCTGTGTGCTCCAAATAATC | 480 |
| QY | 481 | ATGATGTAYATNNTYNTGTNTTYTYTNAACNTYNTGTYNTYATNTAGABATGATNTAYTYG | 540 |
| Db | 481 | ATGATGTATCACTTCTGTGTCTTCTCTCACTTGTGTGCTCATGTGAATGATATATATTGC | 540 |
| QY | 541 | TAYMNAARGTNMSNARCGNGARGARCGNCGNCARGARPAAYGCNWSNGAYTAYTYTNGN | 600 |
| Db | 541 | TACAGAAAGGCTCTCAAAAACGGAAGAGGACAGCCCAAGAAAACGCTGTGACTACCTTGCC | 600 |
| QY | 601 | ATTNCCMSNGARAAYAAARGARPAAYWSNCCNGNTNCCNNGNARGAR | 645 |
| Db | 601 | ATCCCATCTGAGAACMAAGAAATCTCTCGGTGATCCAGTGTGAGGAA | 645 |

| | | | | | |
|------------|---|-------------|------|--------|-----------------|
| RESULT 2 | | | | | |
| CR609664 | | | | | |
| LOCUS | CR609664 | 2555 bp | mRNA | linear | HTC 21-JUL-2004 |
| DEFINITION | full-length cDNA clone CSUDF023YA09 of Fetal Brain of Homo sapiens (human) . | | | | |
| ACCESSION | CR609664 | | | | |
| VERSION | CR609664.1 | GI:50490471 | | | |
| KEYWORDS | HTC; CNSLT cDNA. | | | | |
| SOURCE | Homo sapiens (human) | | | | |
| ORGANISM | Homo sapiens | | | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | | | | |
| AUTHORS | Li, W.B., Gruber, C., Jesssee, J. and Polayes, D. | | | | |
| TITLE | Full-length cDNA libraries and normalization | | | | |
| JOURNAL | Unpublished | | | | |
| REMARK | Contact : Feng Liang Email : fliang@fatech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Paradise Avenue | | | | |
| REFERENCE | Genoscope. 2 (bases 1 to 2555) | | | | |
| AUTHORS | Direct Submission | | | | |
| TITLE | Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : | | | | |
| JOURNAL | BP 191 91006 Evry cedex - FRANCE (E-mail : segreif@genoscope.cns.fr - Web : www.genoscope.cns.fr) | | | | |
| COMMENT | 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five primetimed and enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. | | | | |
| FEATURES | Location/Qualifiers | | | | |

```
source
1. .2555
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF023YA09"
/tissue_type="Fetal brain"
/plasmid="pCMVSPORT_6"
```

| | | | | | | |
|----|-----------------------|--|-----------------|-----------|-------------|--|
| | Query Match | 72.2% | Score 466, | D8 B3, | Length 255; | |
| | Best Local Similarity | 58.8% | Pred. No. 4, | de-105; | | |
| | Matches 379; | Conservative 145; | Mismatches 121; | Indels 0; | Gaps 0; | |
| OY | 1 | ATGCCGNCNTTAAVWGNVNTTTCCNYTCNGCMWSNVTNGTNATNTAATYTGGGTINMSN | 60 | | | |
| Db | 235 | ATGCCCTGCCTTCAATAATATTGTTTCCCTCGAGCTTCTCTCGTGCTTAATCTACCTGGTCAGT | 294 | | | |
| OY | 61 | GTNNGYTTYCCNGNTTYGTNGARGTMCNMWSNGARACNARGONGNRCARGNAAYCCN | 120 | | | |
| Db | 295 | GTCGTCTTCCCTGTGTGTGTGAAGTCCCTCGAGACGAGAGCCGTGCAGGGCACC | 354 | | | |
| OY | 121 | ATGAARYTNMGNTGYATNMSNTGYATGAARMNGARGARTNGARGCNACNACNGTNGT | 180 | | | |
| Db | 355 | ATGAAGCTGGCGCTCATCTCGTAGTAAGAAGAGAGAGGTGAGGCAACAAGTGCTG | 414 | | | |
| OY | 181 | GARTGCTTTAYMNCNGARGNGNAABAATTTTTYNATNTAYGARTLYMGNAAVGN | 240 | | | |
| Db | 415 | GAATGGTTCTACAGGCCCCGAGGGGGGTPAAAGATTTCTTATTTACAGATATCGAATGGC | 474 | | | |
| OY | 241 | CAYCARBARGTNGARWMSNCCNTTTCARGMNGVYTNCAITGGAAYAGSMNSNAARGAYTN | 300 | | | |
| Db | 475 | CACAGAGAGGTGAGAGAGACCCCTTTCAAGGGGGCTGTCAGTGGAAATGGCACAGAACCTG | 534 | | | |
| OY | 301 | CARAGYGTNMSNATNACNGTNYTNAAYGTNACNTNAAAYGAYMSNGENTYNTYAACNTG | 360 | | | |
| Db | 535 | CAGACGCTGCCAACAACCTGCTCMAAGTCACTGTGAACACACTGTGGCCCTCACTACCTGC | 594 | | | |
| OY | 361 | AA YGTNMSNMNGARTTYGARITYGARCGNCAVMGNCNTTYYGTNABRACNACMNGVTN | 420 | | | |
| Db | 595 | AATGTGTCCCGGAGATTGTGAGTTTGAGGCCCATTCGGCCCTTTGTGAAGAAGACGGCGCTG | 654 | | | |
| OY | 421 | ATNCCNNTYMNNGTNAACNGARGARGCNGANGARGAYTTVA CNMSNGTNGTMSNGARATN | 480 | | | |
| Db | 655 | ATCCCTCTAAGAGTCAACCGAGAGGCTGGAGAGAGACTTCACTCTGTGGTCTCAGAAATC | 714 | | | |
| OY | 481 | ATGATGTGAAYATNTYNTYNGTNTTYTNAACNTYNTGTYNTNATPNGARATGANTVATGY | 540 | | | |
| Db | 715 | ATGATGTGACATCTCTTGTGGTCTTCTCTCACTGTGTGGTGTCTCATCGAGATGATATATTCG | 774 | | | |
| OY | 541 | TAYVNABRGTMNSNAARGCNGARGARGCNGCNCARGAPAAAYGCMNSNGAYTYAYTNGCN | 600 | | | |
| Db | 775 | TACAGAAAGGTCTCAAAAAGCCGAAGAGGACGACCACAAAGAAACGGGTCTGACTACCTGGC | 834 | | | |
| OY | 601 | ATNCCMNSNGARAAYARAGAPAAVMSNGCNGTNCNGTNGARBAR | 645 | | | |
| Db | 835 | ATCCCATCTGAGAACAGAGAAATCTGCGGTACCACTGGAGGAA | 879 | | | |

| LOCUS | DEFINITION | VERSION | KEYWORDS | SOURCE | ORGANISM | REFERENCE | AUTHORS |
|-----------------------|---|-------------|----------|---------|-----------------|-----------|---------|
| RESULT 3 HSM801563 | HSM801563 | 4052 bp | mRNA | 1linear | HTC 22-SEP-2004 | | |
| LOCUS | Homo sapiens mRNA; cDNA DKFP761F182 (from clone DKFP761F182). | | | | | | |
| DEFINITION | AL136589 | | | | | | |
| VERSION | AL136589.1 | GI:13726680 | | | | | |
| KEYWORDS | HTC. | | | | | | |
| SOURCE | Homo sapiens (human) | | | | | | |
| ORGANISM | Homo sapiens | | | | | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | | | | | | |
| AUTHORS | 1 (bases 1 to 4052) | | | | | | |
| | Ottewaelder, B.; Obermaier, B., Deutschenburg, S., Schaipp, A., Mewes, H.W., Well, B., Amid, C., Oeanger, A., Fobo, G., Han, M. and Wiemann, S. | | | | | | |

CONSRM TITLE JOURNAL

The German cDNA Consortium
Direct Submission
Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764
Neuberberg, GERMANY

COMMENT

Research Center (DFEZ); Email s.wiemann@dfz-heidelberg.de;
sequenced by Medigenomix (Martinsried/Germany) within the cDNA
sequencing consortium of the German Genome Project.
This clone (DKFZp761F182) is available at the RZPD Deutsches
Resourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneid=DKFZp761F182
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.

FEATURES

source

```

1..4052
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFZp761F182"
/db_xref="taxon:9606"
/clone="DKFZp761F182"
/tissue_type="amygdala"
/clone_id="761 (synonym: hamy2). Vector pSPORT1; host
DH10B; sites NotI + SalI"
/dev_stage="adult"
/notes="voltage-gated sodium channel beta-3 subunit"
1..4052
/gene="DKFZp761F182"
804..1451
/gene="DKFZp761F182"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAB6524.1"
/db_xref="GI:13276681"
/db_xref="GOA:Q9NY72"
/translation="MPAPRLPLASLVITYVSCPPVCEVSETEAOGNPMKLR
CISGMKREYBATTVMFPRPBGKPLIYTRNGHOEYSPFQGLQNGSLDLD
VSTVLNVLNDSGLYCNVSRFEFPAHRPVTYRLIPLRVEEGEDPTSVSSI
MNYLIVFLTLMLEIMYICRYKSKAEBAENASDYLAIPSENKENSAPVPEE"

```

ORIGIN

Query Match 72.2%; Score 466; DB 3; Length 4052;

Best Local Similarity 58.8%; Pred. No. 6.1e-105; Indels 0; Gaps 0;

Matches 379; Conservative 145; Mismatches 121;

```

1 ATGCCGNCNTTAAAYMGNYTTCYTCNYTNGCMNSNTNGTNTATNTATYTGCGTMSN 60
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
804 ATGCTGCTTCAATATATTTCTTCCCTGCTTCTCTCGTCTTACTACTGCTGCTAGT 863
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
61 GTTGTGTTTCGNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGT 120
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
864 GTCTGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 923
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
121 ATGAARYTNGNGYATNTMTGATGAARMNGARNGTNGARNGTNGARNGTNGARNGTNG 180
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
924 ATGAAGCTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 983
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
181 GARTGTGTTTAYMNGCNGARNGGNGAARAGATTTTNTATNTATNTATNTATNTATNTAT 240
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
984 GAATGCTTCTACAGGCCCGAGGGGGGCTTAAGATTTCTTATTTACAGATATCGGAATG 1043
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
241 CAYCARGARTNGARMSNCCNTTYCARGMNGYNTNCACTGGAAYGAGNMSNARAGAYTN 300
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1044 CACGAGAGGTGGAGAGCCCTTTCAGGGGGGCTGAGTGGATGAGTGGAGAGAGACTG 1103
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
301 CARGAYGTNGNATNAGCNGTNTTAAAYGTNAANTNAYGNGNGYNTTAAACNGY 360
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1104 CAGAGCTGTTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1163
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
361 AAYGTNSNNGARITTYGARTTYGARGCNCAYMNGCNGTNTTAAARACNACMNGYTN 420
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1164 AATGTGTCCCGGAGATTGAGATTGAGGCGCATGCGCCCTTGTGGAAGACGAGCGGCTG 1223
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Qy 421 ATNCNTNMGNTNACNGARCGNGGNGARGATTTTACNMSNGTNGTNSGADATN 480
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1224 ATCCCCCTAAGAGTCCACCGAGGAGCTGGAGAGACTTCACTTGTGCTCAGAAATC 1283
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 481 ATGATGTAATNTNNTNNTNNTTNTTNTATNTGNTNGTNTNTATNTGATGATNTATGY 540
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1284 ATGATGTAATCTCTTCTGCTTCTTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1343
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 541 TAYMNAARGTNSNNAARCGNGARCGNGCNGCARAAYGCMNSNGAYTAYTNGCN 600
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1344 TAGCAAGAGTCTCAAAAGCGGAGAGAGCCCAAGAAAACGGTCTGACTACTGCTGCC 1403
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 601 ATNCNMSNGARPAAYARARPAAYWSNGCNGTNCNGTNGARGAR 645
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1404 ATCCCATCTGAGAACAGAGAACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1448
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 4
BX420015 1062 bp mRNA linear EST 01-MAY-2004
LOCUS
DEFINITION
BX420015 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
CS0DF023YA09 5-PRIME, mRNA sequence.
ACCESSION
BX420015
VERSION
BX420015.2 GI:46929710
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1062)
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 13, 2003 this sequence version replaced gi:30646738.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 Evry cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr

```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

This sequence belongs to sequence cluster 6147.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?cs=CS0DF023YA050P1&c=6147.r.

FEATURES

source

```

1..1062
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF023YA09"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_id="Homo sapiens FETAL BRAIN"
/notes="Oxgan: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

```

ORIGIN

Query Match 72.2%; Score 465.6; DB 5; Length 1062;

Best Local Similarity 58.6%; Pred. No. 3.9e-105; Indels 0; Gaps 0;

Matches 378; Conservative 146; Mismatches 121;

```

Qy 1 ATGCCGNCNTTAAAYMGNYTTCYTCNYTNGCMNSNTNGTNTATNTATYTGCGTMSN 60
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 234 ATGCTGCTTCAATATATTTCTTCCCTGCTTCTCTCGTCTTACTACTGCTGCTAGT 293
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 61 GTTGTGTTTCGNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGT 120
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db 294 GTCCTGCTCCCTGTGTGTGTGGAAGTGCCTCGAGACGGAGCGCGTGCAGAGGCC 353

Qy 121 ATGAATYTMNGNTGYATMNSNTGYATGAARMNGARAGRTNARGCNAACNAGNTGN 180

Db 354 ATGAAGCTCGCGCTCATCTCTGCATGAAGAGAGAGGTGAGCCACACGGGTGGT 413

Qy 181 GARTGTTTAAVWNCNGARNGNGNAARGATTTTNTATNTAYGARTAYMNAAYGN 240

Db 414 GATGTTTCTACAGCGCCGAGGGCGGTAAAGATTTCTTATTTACAGATGTGGAATGGC 473

Qy 241 CAYCARGATNGARMSNCCNTTYCARGNMGYNTCARTGAAVYAGNSNAARGAYTN 300

Db 474 CACCAAGAGGTGAGAGCCCTTTCAAGGGCGCTCGACGTGAAATGAGCAAGACCTG 533

Qy 301 CARGATYTMNSNATNACNGNTYNTAAVGTNACNTYNTAAVAGVNSNGNTNTAYACNTGY 360

Db 534 CAGGAGGTGTCATCACTGTGCTCAACGTCACTGAAACGACTGTGCTTACACCTGC 593

Qy 361 AAVGTMSMNGNARTTYGARTTYGARGCNCAYMNCNTTYGTNABACNAACNMGYTN 420

Db 594 AATGTCTCCCGAGGAGTTTGAAGTTCGCGCATCGCCCTTTGTGAAGACGACGGCTG 653

Qy 421 ATNCCNYTMNGTNAACNGARAGCNGNGARAGATTYAACMNSNGTNGTMSNGARATN 480

Db 654 ATCCCTTAAGATCAACCGAGAGCTGGAGAGGACTTCACTGTGTGCTCAGAAATC 713

Qy 481 ATGATGTAAVNTYNTNGTNTTYTNTACNTYNTGYTNTATNTAGATNTATYTG 540

Db 714 ATGATGTACATCTTCTGTCTTCTCACTGTGTGCTGCTCAACGATGATATATGTC 773

Qy 541 TAYMNAAGTMSNABRGNGARAGCNGCNGARAGAAVYGCNMSNGATYATYNTGNC 600

Db 774 TACAGAAAGTCTCAAAAGCGAAGAGGAGCCCAAGAAACGCTGACTACCTTGGC 833

Qy 601 ATNCCMWSNGARAAYARAGARAAYMNGCNGTNCNGTNGARGAR 645

Db 834 ATCCATCTGAGAACAGAGAACTCTGCGGTACCGAGTGAGGAA 878

RESULT 5
AY419147 648 bp DNA linear GSS 12-DEC-2003

LOCUS AY419147
DEFINITION Mus musculus HCM6793 gene, VIRUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY419147
VERSION AY419147.1 GI:39775107
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 648)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shtinsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
These sequences were made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source Location/Qualifiers
1..648 /organism="Mus musculus"

gene /mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..648
/locus_tag="HCM6793"

ORIGIN

Query Match 71.5%; Score 461.2; DB 9; Length 648;
Best Local Similarity 58.3%; Pred. No. 3,8e-104;
Matches 376; Conservative 145; Mismatches 124; Indels 0; Gaps 0;

Qy 1 ATGCGNCCNTTYAAVWNTYNTTYCCNTYNGCNSMYNTGNTYNTATNTAYTGGTMSN 60

Db 1 ATGCTGCTCCCTCAACAGATGCTTCCCTAGCTTCTCTAGTGTCTATCTAATGAGTCA 60

Qy 61 GTTGTGTTCCGNTNGTNGTNGARNTCCNMSNGARAACNGAGCNGTNCARGNAAYCCN 120

Db 61 GTCCTGCTCCCTGTGTGTGTGTAAGAACCTCGGAGACGAAAGCCCTGACGGGCAATTC 120

Qy 121 ATGAATYTMNGNTGYATMNSNTGYATGAARMNGARAGRTNARGCNAACNAGNTGN 180

Db 121 ATGAAGCTAGATGATCTCTCTGCATGAAGAGAGAGAGTGAAGCCACACTAGT 180

Qy 181 GARTGTTTAAVWNCNGARNGNGNAARGATTYTNTATNTAYGARTAYMNAAYGN 240

Db 181 GAGTGTCTTACAGCGCTGAGGGCGGTAAAGATTTCTTATATATAGTATCGAAATGGC 240

Qy 241 CAYCARGATNGARMSNCCNTTYCARGNMGYNTCARTGAAVYAGNSNAARGAYTN 300

Db 241 CACCAAGAGGTGAGAGCCCTTTCAAGGGCGCTCGACGTGAAATGAGCAAGACCTG 300

Qy 301 CARGATYTMNSNATNACNGNTYNTAAVGTNACNTYNTAAVAGVNSNGNTNTAYACNTGY 360

Db 301 CAGGAGGTATCATCACTGTGCTCAATGTCACTGTAATGACTGTGCTTACACATGT 360

Qy 361 AAVGTMSMNGNARTTYGARTTYGARGCNCAYMNCNTTYGTNABACNAACNMGYTN 420

Db 361 AATGTCTCAAGGAGTGTGATGATGGAAGCAACCGCCCTTTGTGAAGACCAACAACTA 420

Qy 421 ATNCCNYTMNGTNAACNGARAGCNGNGARAGATTYAACMNSNGTNGTMSNGARATN 480

Db 421 ATACCCCTCGAGATCACTGAAGGGCGGAGAAAGACTTCACTCTGTGTCTGGAATTC 480

Qy 481 ATGATGTAAVNTYNTNGTNTTYTNTACNTYNTGYTNTATNTAGATNTATYTG 540

Db 481 ATGATGTACATCTCTCTGTCTTCTCACTGTGTGCTGTTATAGATGATCTATGGC 540

Qy 541 TAYMNAAGTMSNABRGNGARAGCNGCNGARAGAAVYGCNMSNGATYATYNTGNC 600

Db 541 TACAGAAAGTCTCTAAGGCCGAAAGGAGCCTCAGGAAATGCGTCACTACCTTGGC 600

Qy 601 ATNCCMWSNGARAAYARAGARAAYMNGCNGTNCNGTNGARGAR 645

Db 601 ATCCCTTCAAGAACAGAGAACTCTGTGTATCCCTGTGAGGAA 645

RESULT 6
AK076466 1359 bp mRNA linear HTC 03-APR-2004

LOCUS AK076466
DEFINITION Mus musculus 0 day neonate head cDNA, RIKEN full-length enriched
library, clone:4833414B02 product:VOLTAGE-GATED SODIUM CHANNEL
BETA-3 SUBUNIT (SODIUM CHANNEL BETA 3 SUBUNIT) homolog [Rattus
norvegicus], full insert sequence.
ACCESSION AK076466
VERSION AK076466.1 GI:26345409
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 Carninci,P. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)

JOURNAL MEDLINE
PUBMED 99279253
REFERENCE 10349636

AUTHORS 2

TITLE Carinci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komoto, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.

JOURNAL Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

MEDLINE Genom. Res. 10 (10), 1617-1630 (2000)

PUBMED 20499374

REFERENCE 11042159

AUTHORS 3

TITLE Shibusawa, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carinci, P., Komoto, H., Akiyama, J., Nishi, K., Kitasuna, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kasai, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kiru, A., and Hayashizaki, Y.

JOURNAL RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer

MEDLINE Genome Res. 10 (11), 1757-1771 (2000)

PUBMED 20530913

REFERENCE 11076861

AUTHORS 4

TITLE THE RIKEN GENOME EXPLORATION RESEARCH GROUP PHASE II TEAM AND THE PANTOM CONSORTIUM.

JOURNAL Functional annotation of a full-length mouse cDNA collection

MEDLINE Nature 409, 685-690 (2001)

PUBMED 11076861

REFERENCE 5

AUTHORS THE PANTOM CONSORTIUM AND THE RIKEN GENOME EXPLORATION RESEARCH GROUP PHASE I & II TEAMS.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

MEDLINE 6 (bases 1 to 1359)

PUBMED Adachi, J., Aizawa, K., Akabira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Bono, H., Carinci, P., Fukuda, S., Fukuishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoka, T., Horii, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kaubukawa, T., Kato, H., Kawai, J., Kojima, Y., Komoto, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toyota, T., Yamamura, T., Yamana, I., Yasuniishi, A., Yoshida, K., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.

JOURNAL Direct Submission

MEDLINE Submitted (16-MAR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Encyclopedic Project of Genome Exploration Research Group in Riken

PUBMED Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

REFERENCE Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/.
Location/Qualifiers
1..1359
/organism="Mus musculus"
/mol_type="RNA"
/strain="C57BL/6J"
/db_xref="PANTOM_DB:4833414B02"
/cd_xref="caxon:10090"
/clone="4833414B02"
/issue_type="head"
/clone_lib="RIKEN full-length enriched mouse cDNA library"

[illegible]

VERSION AK049747.1 GI:26340475
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE AUTHORS Carinci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE AUTHORS Carinci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
TITLE Itoh, M., Komoto, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
JOURNAL Normalization and subtraction of cap-trapper-selected cDNAs to
MEDLINE prepare full-length cDNA libraries for rapid discovery of new genes
PUBMED Genome Res. 10 (10), 1617-1630 (2000)
11042159
REFERENCE AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carinci, P.,
TITLE Komoto, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
JOURNAL Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
MEDLINE Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
PUBMED Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matshiki, M.,
11076861 Okazaki, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
TITLE Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
JOURNAL RIKEN integrated sequence analysis (RISA) system-384-format
MEDLINE sequencing pipeline with 384 multicapillary sequencer
PUBMED Genome Res. 10 (11), 1757-1771 (2000)
20530913
REFERENCE AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
TITLE FANTOM Consortium.
JOURNAL Functional annotation of a full-length mouse cDNA collection
MEDLINE Nature 409, 685-690 (2001)
PUBMED 11076861
REFERENCE AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research
TITLE Group Phase I & II Team.
JOURNAL Analysis of the mouse transcriptome based on functional annotation
MEDLINE of 60,770 full-length cDNAs
PUBMED Nature 420, 563-573 (2002)
6 (bases 1 to 3549)
REFERENCE AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carinci, P.,
TITLE Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
JOURNAL Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
MEDLINE Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
PUBMED Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
11076861 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
TITLE Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohnato, N.,
JOURNAL Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
MEDLINE Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
PUBMED Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
11076861 Takada, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
TITLE Muramatsu, M., and Hayashizaki, Y.
JOURNAL Direct Submission
MEDLINE Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
PUBMED Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
COMMENT CDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in Riken.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/.

FEATURES
source Location/Qualifiers
1..3549
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM DB:C530046B12"
/dev_stage="12 days embryo"
/clone="C530046B12"
/issue_type="spinal cord"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="12 days embryo"
232..879
/note="unlabeled protein product; VOLTAGE-GATED SODIUM
CHANNEL, BETA-3 SUBUNIT (SODIUM CHANNEL BETA 3 SUBUNIT)
homolog [Rattus norvegicus] (SPR1Q9J00, evidence: PASTY,
95.5%ID, 100%length, match=646)
putative"
/codon_start=1
/protein_id="BAC33901.1"
/db_xref="GI:26340476"
/translation="MPAPRRLPLASLVLIYWRVCPVCPVSEFTEAVQNSMELR
CISCKRESVENVTVEMFYRPEGGQFLITERYKHQVSEFPQRLQNSKIDQ
VITVTLNVLNSGLYTCNVSRSEFEFEARPPVKTTRLPLRVTEAGDFTSVSEI
MMYILVFLTLWLFLEMICYRKVSKABEBAQENASDYLAIPESENKENVPEE"

ORIGIN
Query Match 71.5%; Score 461.2; DB 3; Length 3549;
Best Local Similarity 58.3%; Pred. No. 9.1e-104;
Matches 376; Conservative 145; Mismatches 124; Indels 0; Gaps 0;
1
ATGCCGCGCTTAAAYGANTTTCCNTYNGCNSNYTNGTNTAATNTAATGCGTWSN 60
232 ATGCCGCGCTTAAAGAGATTGCTCCCTAGCTTCTAGTGTCTATCTACTGGTCAGA 291
61 GNTGTGTTTCGNTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 120
292 GTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 351
121 ATGAATYTMGNTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 180
352 ATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 411
181 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 240
412 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 471
241 CAYCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
472 CACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 531
301 CARGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
532 CAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 591
472 CACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 531
301 CARGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 360
532 CAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 591
592 AATGTCACAGGAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 651
421 ATNCCNTYTMGNTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
652 ATNCCNTYTMGNTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 711
481 ATNCCNTYTMGNTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
712 ATNCCNTYTMGNTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 771
541 TAYMGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
772 TACAGAAAGTCTCTAAGCCGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 831
601 ATNCCNTYTMGNTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 645
832 ATNCCNTYTMGNTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 876

| RESULT 8 | 4149 bp | mRNA | linear | HTC 03-APR-2004 |
|------------|---|-------------|--------|-----------------|
| AK049286 | | | | |
| LOCUS | | | | |
| DEFINITION | AK049286 | | | |
| | Mus musculus ES cells cDNA, RIKEN full-length enriched library, clone:G330019103 product:VOLTAGE-GATED SODIUM CHANNEL BETA-3 SUBUNIT (SODIUM CHANNEL BETA 3 SUBUNIT) homolog [Rattus norvegicus], full insert sequence. | | | |
| ACCESSION | AK049286 | | | |
| VERSION | AK049286.1 | GI:26093400 | | |
| KEYWORDS | HTC, CAP trapper. | | | |
| SOURCE | Mus musculus (house mouse) | | | |
| ORGANISM | Mus musculus | | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | | |
| AUTHORS | 1 Carninci, P. and Hayashizaki, Y. | | | |
| TITLE | High-efficiency full-length cDNA cloning | | | |
| JOURNAL | Meth. Enzymol. 303, 19-44 (1999) | | | |
| MEDLINE | 99279253 | | | |
| PUBMED | 10349636 | | | |
| REFERENCE | 2 | | | |
| AUTHORS | Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komano, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. | | | |
| TITLE | Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes | | | |
| JOURNAL | Genome Res. 10 (10), 1617-1630 (2000) | | | |
| MEDLINE | 20499374 | | | |
| PUBMED | 11042159 | | | |
| REFERENCE | 3 | | | |
| AUTHORS | Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komano, H., Akiyama, J., Nishi, K., Kitasuna, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. | | | |
| TITLE | RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer | | | |
| JOURNAL | Genome Res. 10 (11), 1757-1771 (2000) | | | |
| MEDLINE | 20530913 | | | |
| PUBMED | 11076861 | | | |
| REFERENCE | 4 | | | |
| AUTHORS | The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. | | | |
| TITLE | Functional annotation of a full-length mouse cDNA collection | | | |
| JOURNAL | Nature 409, 685-690 (2001) | | | |
| MEDLINE | 5 | | | |
| PUBMED | | | | |
| REFERENCE | | | | |
| AUTHORS | | | | |
| TITLE | The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. | | | |
| JOURNAL | Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs | | | |
| MEDLINE | Nature 420, 563-573 (2002) | | | |
| PUBMED | 6 (bases 1 to 4149) | | | |
| REFERENCE | | | | |
| AUTHORS | Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hori, F., Imocani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Komano, H., Kouda, M., Koya, S., Kurahara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Nunazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takanashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y. | | | |
| SOURCE | Muramatsu, M. and Hayashizaki, Y. | | | |
| ORGANISM | Direct Submission | | | |
| REFERENCE | Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail:genome-res@gs.c.riken.jp]. | | | |

[illegible]

```

Db      950 CTACAGAAAGCTCTCTAAGCCCAAGAGCAGCTCAGAAATGCTGTGACTTCTGC 1009
Qy      600 NATCCNWSNGARAAAYARAGARAAVWSNGCNGTNGCNGTNGAR 645
      1010 TATCCCTTCAGAGACAGAGAACTGTGTGATCCCGTGAGAGAA 1055

RESULT 9
BM933157      636 bp      mRNA      linear      EST 13-MAR-2002
LOCUS      BM933157
DEFINITION      UI-M-BH3-beg-d-04-0-UI.r1 NIH BMAP M S4 Mus musculus cDNA clone
ACCESSION      BM933157
VERSION      BM933157
KEYWORDS      BM933157.1 GI:19392309
SOURCE      EST.
ORGANISM      Mus musculus (house mouse)
MUSCULUS      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 636)
Normalizaion and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
JOURNAL      97044477
MEDLINE      8889548
COMMENT      Contact: Chin, H
      National Institute of Mental Health
      6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
      20892-9643, USA
      Tel: 301 443 1706
      Fax: 301 443 9890
      Email: mestr@mail.nih.gov
      cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
      cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
      DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
      Clone Distribution: Researchers may obtain clones from Research
      Genetics (www.resgen.com).
      Seq primer: M13 REVERSE
      Location/Qualifiers
      1. 636
      /organism="Mus musculus"
      /mol_type="mRNA"
      /strain="C57BL/6J"
      /db_xref="taxon:10090"
      /clone="UI-M-BH3-beg-d-04-0-UI"
      /dev_stage="27-32 days"
      /lab_host="DH10B (Life Technologies)"
      /note="Vector: PT73D-Pac (Pharmacia) with a modified
      polylinker; Site 1: Not I; Site 2: Eco RI; The
      NIH_BMAP_M_S4 library is a subtracted library of a series,
      ultimately derived from a mixture of individually tagged
      normalized libraries from ten regions of the mouse brain
      (cerebellum, brain stems, olfactory bulbs, hypothalamus,
      cortex, amygdala, basal ganglia, pineal gland, striatum,
      hippocampus) after a series of subtractions to reduce the
      representation of cDNAs from which ESTs had already been
      generated. The following serially subtracted libraries
      were generated in this process: NIH_BMAP_M_S3.1,
      NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, NIH_BMAP_M_S4,
      NIH_BMAP_M_S2, NIH_BMAP_M_S1. The subtracted library
      (NIH_BMAP_M_S4) was constructed as follows: PCR amplified
      cDNA inserts from NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and
      NIH_BMAP_M_S3.1 clones from which 3' ESTs had been derived
      was used as a driver in a hybridization with a pool of
      the NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and NIH_BMAP_M_S3.1
      libraries in the form of single-stranded circles. The
      remaining single-stranded circles (subtracted library)
      was purified by hydroxyapatite column chromatography,
      converted to double-stranded circles and electroporated
      into DH10B bacteria (Life Technologies) to generate the
  
```

ORIGIN

Query Match 69.4%; Score 447.8; DB 5; Length 636;
 Best Local Similarity 58.1%; Pred. No. 8.5e-101;
 Matches 365; Conservative 142; Mismatches 121; Indels 0; Gaps 0;

NIH_BMAP_M_S4 library. This procedure has been previously
 described (Bonaldo, Lennon and Soares, Genome Research
 6:791-806, 1996)"

```

Qy      1 ATCCNCGNTTYAAYMGNTTNTTYCCNYTNGCNWSNYTNGTNTAATNTAYTGGTMSN 60
      7 ATGCCCTGCTTCAACAGATTTGCTCCCTTACGCTTCTAGTGTCTATCTACATGGGTCAG 66
Db      61 GTNTGYTTCGCGTNGTNGTNGTNGTNGCNWSNGARCGNCTNGARCGNAAYCCN 120
      67 GTCTGCTTCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 126
Qy      121 ATGAARYTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGTNGT 180
      127 ATGAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 186
Db      181 GARTGTTTATMGNCNGARNGNGNARAGATTTTNTAINTATYARTATMGNAAYG 240
      187 GAGTGTTCTACAGGCGCTGAGGCGGTAAAGATTCTTATATATGATGATAAATGCG 246
Qy      241 CAYCARGARTNGARBSNCCNTTYCARGNMGNNTNCTARTGGAAYGNSNARAGAYTN 300
      247 CACGAGAGGTGAGAGGCCCCCTTCAGAGTCTGTCTCAGAGTGTGTGTGTGTGTGTGTGT 306
Db      301 CARGAYTMSNATNACNGTNTNTAAYGTNACNYTNAAYGAVWSNGNTNTATYACNTGY 360
      307 CAGGACGTATCATCATCTGTCTCAATGTCACTCTGAATGATCTGCGCTCTACATG 366
Qy      361 AAYGTWSMNGARTTYGARTTYGARGCNCAYMGNCNTTYGTNAARACNMGNTYN 420
      367 AATGTCTCAGGAGATTTAGATTCAGAGACACCGGCTTGTGTGAAGACCAAGACTA 426
Db      421 ATNCCNYTMNGTNCNGARCGNCGNGNGARAGATTTTACNWSNCTNGTNGSNGARATN 480
      427 ATACCCCTCGAGTCACTGAAGAGGCGGAGAGAACTTCACTCCCTGTGTGTGGAATC 486
Qy      481 ATGATGATAYATNTYNTGNTNTTYTNAACNTNTGTNTYNTAATGATATATYTG 540
      487 ATGATGATATCTCTCTGCTCTTCCCTCACTTGTGTGTGTGTGTGTGTGTGTGTGTGT 546
Db      541 TATMGTAARTGTMNSNARCGNARGCGNCGNCAAGAAAGCNSNAGATATYTNCGN 600
      547 TACAGAAAGTCTCTAAGGTCTGAAGAGGAGCTCAGAAATGCTGTGATCTACTTGTCT 606
Qy      601 ATNCCNWSNGARAAAYARAGARAAVWSNG 628
      607 ATCCCTTCAGAGACAGAGAACTGTG 634

RESULT 10
BX445002      975 bp      mRNA      linear      EST 04-MAY-2004
LOCUS      BX445002
DEFINITION      CS0DF023YA09 5-PRIME, mRNA sequence.
ACCESSION      BX445002
VERSION      BX445002
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 975)
REFERENCE      Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
      Full-length cDNA libraries and normalization
      Unpublished (2001)
      On May 15, 2003 this sequence version replaced gi:30782286.
      Contact: Genoscope
      Genoscope - Centre National de Sequencage
  
```


| | | | |
|------------|---|--|--|
| Qy | 1 | ATGCCGNGNTTAAATMGATNTTTCYCNATNGGCMNSNTNGTNTAATNTAATGATGGGTMMNS | 60 |
| Db | 356 | ATGCCCTGCTTCAACAGATGCTTCCCTCCTAGTCTCTAGTGCTCATCTACCTAGGG--- | A 411 |
| Qy | 61 | GTNTGTYTTCNGTNTGYGTNGARGTNCNMSNGARACNGARCGNGTNCARGSNAAYCCN | 120 |
| Db | 412 | GTCTGCTTCCCTGTGTGTAGTAAGTACCCCTGGAGACAGAACCGGTGAGGGCAATTCC | 471 |
| Qy | 121 | ATGAARYYTMMGTGYATMWSNTGYATGAARMNGARGARGTNGARGCNAACNGTNGTN | 180 |
| Db | 472 | ATGAAGCTGAGATGATCTCCTCCGATGAAGAGGAGGAGGTGAGGCCACCTGTAGTG | 531 |
| Qy | 181 | GARTGGTYYTAYMGCCNGARGGNGSNARGAATYYTNTAATNTAATGABTAYVWNAAYGN | 240 |
| Db | 532 | GAGTGTCTTACAGCCCTGAGGGCGGTAAAGATTCCTTATATATAGATATGAAATGGC | 591 |
| Qy | 241 | CAYCARGARGTNGARWNSNCNTTYCARGGNMGNNTNCARTGAAAYGWMNSNARGAAYTN | 300 |
| Db | 592 | CACCAAGAGGTGAGAGAGGCCCTTCCAAAGGTGTGTGAGTGAATGGAGCA--GACCTG | 650 |
| Qy | 301 | CARGAYGTNWSNATNACNGTNTNAAAYGTNACNTNAAAYGAWNSGNATNTAATCNTGY | 360 |
| Db | 651 | CAGGACGTATCCATCACTACTGTCTCATGTCACTCGAATGACTCTGGCGCTCACACATGT | 710 |
| Qy | 361 | AAVGNWSNMNGAATTTYGARTTYTGARGGNCAYMGNCCNTTYGTAAABACNACNNGNTN | 420 |
| Db | 711 | AATGTGTCCAGGAGATTTGATGTTGAAAGCACACGGCCCTTTGTAAAGCACAAAGACTA | 770 |
| Qy | 421 | ATNCCNYYTMNGTNAACNGARGARCGNGNGARGAYTTYACNWSNGTNGTMSNGARATN | 480 |
| Db | 771 | ATACCCCTGCGAGTCACTGAAGAGCGGGAAGAAGACTTCACTCGGTGTCTCGGAAATC | 830 |
| Qy | 481 | ATGATGTATATNTYNTNGTNTTYTNAACNTYNTGTYNTYNTAATNGARATGATNTAYTGY | 540 |
| Db | 831 | ATGATGTATCACTCTCTGCTCTCTCACTCTTGCTGTTATGAGATGATCTATATGTC | 890 |
| Qy | 541 | TAYMNAARGTMSNAAARGCNGARGARCGNGCNCARGAAYVGCNWSNGAYTAYTNGCN | 600 |
| Db | 891 | TACGAAAGGTCTCTTAAGCCCGAAGAGGAGCTTCAAGAAATAGCGTCTGACTACTGTCT | 950 |
| Qy | 601 | ATNCCMWSNGARAAAYARGARAAVWSNCGNCGTNCNCGTNGARGAR 645 | |
| Db | 951 | ATCCCTTCAGAGAACAGAGAGAACTCTGTGTGTAACCCGTGGAGGAA 995 | |
| RESULT 13 | | | |
| BG294174 | | 950 bp | mRNA linear |
| BG294174 | | | EST 21-FEB-2001 |
| LOCUS | 602391245P1 | NIH_MGC_94 | Mus musculus cDNA clone IMAGE:4503250 5' |
| DEFINITION | mRNA sequence. | | |
| ACCESSION | BG294174 | | |
| VERSION | BG294174.1 | | GI:13054543 |
| KEYWORDS | EST. | | |
| SOURCE | Mus musculus | | |
| ORGANISM | Mus musculus (house mouse) | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | |
| AUTHORS | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | |
| TITLE | NIH-MGC http://mgc.nci.nih.gov/ . | | |
| JOURNAL | National Institutes of Health, Mammalian Gene Collection (MGC) | | |
| COMMENT | Unpublished (1999) | | |
| | Contact: Robert Strausberg, Ph.D. | | |
| | Email: gsa@b6-remail.nih.gov | | |
| | Tissue Procurement: The Cepko Laboratory | | |
| | cDNA Library Preparation: Life Technologies, Inc. | | |
| | cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) | | |
| | DNA Sequencing by: InCyte Genomics, Inc. | | |
| | Clone distribution: MGC clone distribution information can be | | |
| | found through the I.M.A.G.E. Consortium/LNL at: | | |
| | http://image.lnl.gov | | |

```

/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:4503250"
/cisue_type="retina"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_94"
/note="Organ: eye; Vector: pCMV-Sport6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 3.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

```

| | | | | |
|-----------------------|-------------------|--------------------|-----------|-------------|
| Query Match | 62.8%; | Score 405.2; | DB 4; | Length 950; |
| Best Local Similarity | 58.4%; | Pred. No. 4.9e-90; | | |
| Matches 348; | Conservative 134; | Mismatches 112; | Indels 2; | Gaps 2; |

| | | | |
|----|-----|---|-----|
| Qy | 1 | ATGCNCCNTTAAVYMGNNTTTCNVTNGCNSMNTNGNTNNATNTATVGGGTMSN | 60 |
| Db | 87 | ATGCTCCCTCTCAACAATGCTTCCCTCAGCTTCTTAAGTCTCATCTACTCTAGGTCA | 146 |
| Qy | 61 | GTNTGYTTTCNGTNTGTGTNGAGTTCNMCMSNBARACNGARGCNGTNCARGNAACN | 120 |
| Db | 147 | GTCTGCTTCCCTGTGTGTGTGAAGTACCTCGAGACAGAAGCGGTGCAAGGCATTTCC | 206 |
| Qy | 121 | ATGAATYTNMGNTGYATNMSNTGYATGAARWNGARGARGTNGARGCNACNAGTNGTN | 180 |
| Db | 207 | ATGAAGCTGAGATGCATCTCCTGCATGAAGGAGGAGGTGAGGCGCACCATCTAGTG | 266 |
| Qy | 181 | GARGGTYYTAYMNCNCGARGGNGNBARGYTTTNTATNTATYGTATMGNAAVGN | 240 |
| Db | 267 | GAGTGTCTTCAACGGCTGAGGGGGGTPAAGATTTCTTAATATYGTATCGAATGCG | 326 |
| Qy | 241 | CAYCARARGTNGARMSNCCNTTTCARCGMNGNTNCARTGGAAYGGMWSNBARGYTN | 300 |
| Db | 327 | CACAGAGAGGTGAGAGAGCCCTTCCAAAGTGTGTGCAGTGGAATGGGAGCAAAAGCCTG | 386 |
| Qy | 301 | CARGAYGTMSNATNACNGTNYTNAAYGTNACNTNAAVGAYSNGNVTNTAYACNTGY | 360 |
| Db | 387 | CAGACCTATTCATCACTGTTCTCAATGTCACTGTGAATGATCTCGACCTTACACAGTGT | 446 |
| Qy | 361 | AAVGTMSNMNGARTTYGARTTYGARGCNCAYMNGCNCNTTGTNANBARCNAACMGVTN | 420 |
| Db | 447 | AAATGTTCAG-GAGTTTGAGTTCGAAGACAACCGGCCCTTTGTGAAGACCAAGAAGCTA | 505 |
| Qy | 421 | ATNCCNTNMGNGTNAACNGARGARGCNGNGARGAYTTVACMSNGTNGTWSNGBATN | 480 |
| Db | 506 | ATACCCCTCGAGTCACTGAAGAAGCGGGAGAAAGACTTCACTCCGTGTGTCCGGAATC | 565 |
| Qy | 481 | ATGATGTAAVNTYNTYNTGTNTTYYTNAQVNTTGYTNTYTNATNGARPATCATNTGY | 540 |
| Db | 566 | ATGATGTAACTCCCTCGTCTTCTCCTCACTTGCGCTGTTTATYTGAGATCATATATGC | 625 |
| Qy | 541 | TAYMGNABGTMSNARGCNGARGCNGCNGCNCARBARAAVGMNSGAVTAYT | 596 |
| Db | 626 | TACGAAAGGTCTTAAAGCCGAAGAGGACAG-TAGGAAATATGCTCTGACTACTT | 680 |

| | |
|------------|--|
| RESULT 14 | |
| BP200910 | |
| LOCUS | 582 bp mRNA linear EST 14-SEP-2004 |
| DEFINITION | BP200910 Sugano cDNA library, amygdala Homo sapiens CDNA clone |
| | AKR03894, mRNA sequence. |
| ACCESSION | BP200910 |
| VERSION | BP200910.1 GI:52050356 |
| KEYWORDS | EST. |
| SOURCE | Homo sapiens (human) |
| ORGANISM | Homo sapiens |

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 582)
 Suzuki,Y., Yamashita,R., Shirota,M., Sakakibara,Y., Chiba,J.,
 Mizushima-Sugano,J., Nakai,K. and Sugano,S.
 Sequence comparison of human and mouse genes reveals a homologous
 block structure in the promoter regions
 Genome Res. 14 (9), 1711-1718 (2004)
 CONTACT: Yutaka Suzuki
 Department of Virology
 Institute of Medical Science, University of Tokyo
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
 Email: yusuzuki@ims.u-tokyo.ac.jp.
 Location/Qualifiers
 1..582
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="AMR03894"
 /issue_type="amygdala"
 /clone_lib="Sugano cDNA library, amygdala"
 ORIGIN
 Query Match 61.2%; Score 395; DB 5; Length 582;
 Best Local Similarity 58.8%; Pred. No. 1.4e-87;
 Matches 323; Conservative 121; Mismatches 105; Indels 0; Gaps 0;
 QY 3 GCCNGCCTTAAATGATTTTCCNTTNGCNSNTYNTATNTATYGGTMSNGT 62
 DB 33 GCGCTGCTTCAATGATGTTTCCCTGCTGCTCTCTATCTACTGAGGTCAGTGT 92
 QY 63 NTGTYTCCNGTGTGTGTGARGTNCNSNGARAGCNGTCCARGNAAYCNAT 122
 DB 93 CTGCTTCCCTGCTGTGTGTGAAGTGCCTGGAGACGAGGCCCTGACAGGCCAAT 152
 QY 123 GAATYTNMGNTGYATYTNMTGYATGARMNGARGATNGARGCNAACNCTNGTNGA 182
 DB 153 GAAGTGCCTGATCTCTGATGAGAGAGAGAGAGAGTGAAGCCACGAGTGTGA 212
 QY 183 RTGCTTATYMGCCNGARGGNGNARAGATTTTNTATNTATYAGATYMGNAAYGNC 242
 DB 213 ATGGTCTACAGGCCCGAGAGCGGTAAAGATTTCTTATTTACAGATATCGAATG 272
 QY 243 YCARGARGNGARWNSCNTTYTCARGNMGNNTNCACTGGAAGWMSNARAGATYTNCA 302
 DB 273 CCAAGAGGTGAGAGGCCCTTTCAGAGGGGCCCTGCAATGCAAGCAAGCACTGCA 332
 QY 303 RGAATGNTSNATNACNGTNTYTNAAAGTNAACNTYTNAAAGATYTNAAAGTNA 362
 DB 333 GACGCTGCATCATGTGTGCTCAACGTCACTGGAAGACTGTGAGCTGACACTGCA 392
 QY 363 YGTNWMNMGATTTGATTTGARGCNCAYMNCNTTYGTNAACNACNMGNTNAT 422
 DB 393 TGTGTCCTCGGATTTGAGTTGAGGCGATCGCCCTTGTGAACACGCGGGTGTAT 452
 QY 423 NCCNTYTNMGNTNACNGARGCNGNGARGATTTTCACWMSNGTNGWMSNARATNAT 482
 DB 453 CCCCCCAAGATGACCGAGAGAGCTGAGAGAGCACTTCACTCTGTGTCTCAGAAATCAT 512
 QY 483 GATGTAATNTYNTGNTTNTYTNACNTYTNAGTNTYTNATNGARATGATNTATYGYTA 542
 DB 513 GATGTACATCTCTGTGTCTCTCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 572
 QY 543 YMGNAARGT 551
 DB 573 CAGAAAGGT 581
 RESULT 15
 CA749311 723 bp mRNA linear EST 09-JUL-2003
 LOCUS UI-M-FY0-cdd-k-17-0-UI-r1 NIH_BMAP_FY0 Mus musculus cDNA clone
 DEFINITION IMAGE:6831594 5', mRNA sequence.

ACCESSION CA749311
 VERSION CA749311.1 GI:25570984
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 723)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA library arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)
 Seg primer: PYX-5.
 Location/Qualifiers
 1..723
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:6831594"
 /issue_type="whole brain"
 /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH_BMAP_FY0"
 /note="Organ: Brain; Vector: pyX-Aac; Site_1: Ecor I;
 Site_2: Not I; The library was constructed according
 to Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with Ecor I adaptor, digested with Not I and then cloned
 directionally into pyX-Aac vector. The library tag
 sequence located between the Not I site and the polyA tail
 is ACCGAGACG. This library was created for the University
 Iowa Brain Anatomy Project (BMAP): "Gene Discovery in the
 Developing Mouse Nervous System", supported by National
 Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
 program coordinator."
 ORIGIN
 Query Match 59.7%; Score 385; DB 6; Length 723;
 Best Local Similarity 58.5%; Pred. No. 4.8e-85;
 Matches 324; Conservative 122; Mismatches 107; Indels 1; Gaps 1;
 QY 1 ATGCCGNCCTTAAATGATTTTCCNTTNGCNSNTYNTATNTATYGGTMSNGT 60
 DB 170 ATGCTGCTTCAACAGATTTGCTTCCCTGAGCTTCTATGTGTCTATCTAGTGGTCA 229
 QY 61 GTNTGYTTCNGTNTGTGTGARGTNCNSNGARAGCNGTNCARGNAAYCN 120
 DB 230 GTCTGCTTCCCTGCTGTGTGATGAGTACCTCGAGACAGAGCCGTGACAGGCAATTCC 289
 QY 121 ATGAATYTNMGNTGYATYTNMTGYATGARMNGARGATNGARGCNAACNCTNGTNG 180
 DB 290 ATGAAGCTGAGATGATCTCTGATGAGAGAGAGAGAGTGAAGCCACTGATGTG 349
 QY 181 GATGCTTATYMGCCNGARGGNGNARAGATTTTNTATNTATYAGATYMGNAAYGNC 240
 DB 350 GAGTGTCTTACAGCGCTGAGGCGGTAAAGATTTCTTATNTATGAGATGCAATGCG 409
 QY 241 CAYCARGARGTNGARWNSCNTTYTCARGNMGNNTNCACTGGAAGWMSNARAGATYTN 300


```

Db      410 CACCAGAGGTGGAGAGCCCTTCCAAAGTCTGCTGCAATGGAGTGGAGCAAGACTG 469
Qy      301 CAGAGYGTNNSNATNACNNTNNAAYGTNACNNTNNAAYGAYWSNGNNTNNAAYACNTGY 360
Db      470 CAGGACGTATCCATCACTGTTCTCAATGTCACCTGAATGACCTGCGCTCTACACATGT 529
Qy      361 AAYGTNWSNNGNARTTYGARTTYGARGCNCAYMGNCCNTTYGTNARACNACMGNNTN 420
Db      530 AATGTGTCAGGAGGATTGAGTTGAGACACACCGGCCCTTGTGAAGACCACAAAGACTA 589
Qy      421 ATNCCNNTNNGNNGTNAACNGARG-ARGCNGNGARGAYTYTTCNWSNGTNGTNSNGARAT 479
Db      590 ATACCCCTGCGAGTCACTGAAGAAGGCGGAGAGACTTCACCTCCGTGCTGCGAAT 649
Qy      480 NATGATGTAYATNNTNNTGNTTNTTYTNAACNNTNNTGTYNTNATNATGATNTNAYTG 539
Db      650 CATGATGTACATCTCTGCTCTCTCACCTTGTGCTGCTGTTATGAGATGATCTATTTG 709
Qy      540 YTAVMGNARGTNW 553
Db      710 CTACAGAAAGTCT 723

```

Search completed: April 7, 2005, 23:29:10
 Job time : 3113 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 7, 2005, 12:50:15 ; Search time 3038 Seconds
(without alignments)
3429.187 Million cell updates/sec

Title: US-09-977-579-2

Perfect score: 1 MPANRRLPLASLVLYWVS.....SDYLAISEKENSAPVEE 215

Sequence:

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q/cgna 1/USPTO.spool_p/US0997579/runat.06042005.163456.23592/app.query.fasta_1.391
-DB=GenEmbl -GFWT=fastap -SUFFIX=ice -MINMATCH=0.1 -LOOPCT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR_SCORE=pcp -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFWT=pcp -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USBR=US0997579.0CGN 1 1 5600 @runat.06042005.163456.23592 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEOUTERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database: GenEmbl:*

1: gb_da:*\n2: gb_hg:*\n3: gb_in:*\n4: gb_on:*\n5: gb_ov:*\n6: gb_pat:*\n7: gb_ph:*\n8: gb_pl:*\n9: gb_pr:*\n10: gb_ro:*\n11: gb_sts:*\n12: gb_sy:*\n13: gb_un:*\n14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match Length | ID | Description |
|------------|-------|--------------------|------|----------------------|
| 1 | 1124 | 100.0 | 1261 | 6 AR359850 Sequence |
| 2 | 1124 | 100.0 | 1261 | 6 AX039100 Sequence |
| 3 | 1124 | 100.0 | 1261 | 9 HSA243396 |
| 4 | 1124 | 100.0 | 5306 | 9 AB032984 Homo sapi |

not pair

| | | | | |
|----|-------|------|--------|--------------|
| 5 | 1116 | 99.3 | 3296 | 9 AB097521 |
| 6 | 1105 | 98.3 | 645 | 6 AX048004 |
| 7 | 1105 | 98.3 | 2220 | 6 AR359849 |
| 8 | 1105 | 98.3 | 2220 | 6 AX039099 |
| 9 | 1105 | 98.3 | 2220 | 6 RNO243395 |
| 10 | 1105 | 98.3 | 2632 | 6 AX048005 |
| 11 | 1105 | 98.3 | 3107 | 10 AF378093 |
| 12 | 1105 | 98.3 | 3108 | 6 AX047984 |
| 13 | 1105 | 98.3 | 3910 | 6 BC070899 |
| 14 | 1097 | 97.6 | 4025 | 10 AK173115 |
| 15 | 1097 | 97.6 | 4169 | 10 BC053919 |
| 16 | 1097 | 97.6 | 4176 | 10 BC058636 |
| 17 | 1091 | 97.1 | 670 | 10 AY049036 |
| 18 | 1028 | 91.5 | 606 | 6 CQ728741 |
| 19 | 857 | 76.2 | 922 | 5 BX935227 |
| 20 | 755 | 67.2 | 912 | 5 BX931708 |
| 21 | 742 | 66.0 | 471 | 6 BD059018 |
| 22 | 714 | 63.5 | 1654 | 5 BC077295 |
| 23 | 477 | 42.4 | 657 | 4 OC035382 |
| 24 | 477 | 42.4 | 657 | 6 AX048006 |
| 25 | 473 | 42.1 | 1335 | 9 HUMVSC1B |
| 26 | 473 | 42.1 | 1404 | 9 HUMSCN1B |
| 27 | 473 | 42.1 | 1525 | 9 BC067122 |
| 28 | 472 | 42.0 | 1490 | 6 AX401978 |
| 29 | 472 | 42.0 | 1490 | 10 RATSCHB1B |
| 30 | 471 | 41.9 | 1557 | 10 BC039140 |
| 31 | 470 | 41.8 | 617 | 6 CQ722293 |
| 32 | 468 | 41.6 | 1125 | 10 MMU46681 |
| 33 | 468 | 41.6 | 1340 | 10 MMU5786 |
| 34 | 468 | 41.6 | 1568 | 10 BC009652 |
| 35 | 423 | 37.6 | 1239 | 9 BC021266 |
| 36 | 400.5 | 35.6 | 140307 | 2 AC135353 |
| 37 | 400.5 | 35.6 | 174285 | 2 AC145484 |
| 38 | 400.5 | 35.6 | 238861 | 2 AC148331 |
| 39 | 400 | 35.6 | 214306 | 2 AC128723 |
| 40 | 396.5 | 35.3 | 127347 | 2 AP000682 |
| 41 | 396.5 | 35.3 | 144833 | 2 AC063921 |
| 42 | 396.5 | 35.3 | 149800 | 2 AC021981 |
| 43 | 396.5 | 35.3 | 172546 | 2 AC024604 |
| 44 | 396.5 | 35.3 | 178169 | 2 AP002765 |
| 45 | 396.5 | 35.3 | 178431 | 2 AP002749 |

ALIGNMENTS

| | | | | | | |
|------------|--|------------------------------------|---------|-----|--------|-----------------|
| RESULT 1 | AR359850 | Sequence 4 from patent US 6593565. | 1261 bp | DNA | Linear | PAT 17-AUG-2003 |
| LOCUS | AR359850 | | | | | |
| DEFINITION | Sequence 4 from patent US 6593565. | | | | | |
| ACCESSION | AR359850 | | | | | |
| VERSION | AR359850.1 | GI:33766660 | | | | |
| KEYWORDS | | | | | | |
| SOURCE | Unknown. | | | | | |
| ORGANISM | Unknown. | | | | | |
| REFERENCE | 1 (bases 1 to 1261) | | | | | |
| AUTHORS | Heslin, P. and Lynam, N. R. | | | | | |
| TITLE | Vehicle interior rearview mirror assembly including an | | | | | |
| JOURNAL | accessory-containing housing | | | | | |
| FEATURES | Patent: US 6593565-A 4 15-JUL-2003; | | | | | |
| SOURCE | Location/Qualifiers | | | | | |
| | 1..1261 | | | | | |
| | /mol_type="genomic DNA" | | | | | |

ORIGIN

Alignment Scores:
Pred. No.: 1.98e-122
Score: 1124.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
Length: 1261
Matches: 215
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-09-977-579-2 (1-215) x AR359850 (1-1261)

QY 1 MetProAlaPheAsnArgLeuPheProLeuAlaSerLeuValLeuIleTyrrTPValSer 20
 Db ATGCGCTGCTTCAATGATGATTTCCCTGGCTTCTCTCGTCTTATCTACTGGGACAGT 435
 QY 21 ValCybPheProValCybValGluValProSerGluThrGluAlaValGlnGlyAsnPro 40
 Db GTCTGCTTCCCTGTGTGTGTGAAGTGCCCTCGAGAGCGAGGCGGTGACAGGACCC 495
 QY 41 MetIleLeuArgCysIleSerCysMetIleAspGluGluValGluAlaThrThrValVal 60
 Db ATGAAGCTGGCGTGCATCTCTCCATGAAAGAGAGAGAGAGGCGCACACGCGTGTG 555
 QY 61 GluTrpPheTyrrArgProGluGlyGlyIleAspPheLeuIleTyrrGluTyrrArgAnGly 80
 Db GAATGGTTCTTACAGGCCCGAGGGCGGTAAAGATTTCTTATTTACGAGTATCGGAATGGC 615
 QY 81 HisGlnGluValGluSerProPheGlnGlyArgLeuGlnTrpAsnGlySerIleAspLeu 100
 Db CACCAAGAGGTGAGAGACCCCTTTCAGGGGCGCTGCGAGTGGACAGACCTG 675
 QY 101 GlnAspValSerIleThrValIleAsnValThrLeuAsnAspSerGlyLeuTyrrThrCys 120
 Db CAGAGCGTGTCCATCACTGTGCTCAACGTCACTGAGACACTCGGCTCTACACCTGC 735
 QY 121 AsnValSerArgGluPheGluPheGluAlaHisArgProPheValIleTyrrThrArgLeu 140
 Db AATGTGTCCCGGAGTTTGAAGTGTGAGCGCATCGCCCTTGTGTGAAGACAGCGCGCTG 795
 QY 141 IleProLeuArgValThrGluGluAlaGlyIleAspPheThrSerValIleSerGluIle 160
 Db ATCCCTTAAGATCACCGAGAGGCTGAGAGACTTCACTCTGTGTCTCAGAAATC 855
 QY 161 MetMetTyrrIleLeuLeuValPheLeuThrLeuTrpLeuLeuIleGluMetIleTyrrCys 180
 Db ATGATGTACATCTCTGTGCTTCTTCAACCTGTGCTGCATCGAGATGATATTTGC 915
 QY 181 TyrArgIleValSerIleAspGluGluAlaHisArgProPheValIleTyrrThrArgLeu 200
 Db TACAGAAAGGTCTCAAAAGCCGAAAGGAGGAGCCCAAGAAACCGCTGACTACCTTGGC 975
 QY 201 IleProSerGluAsnIleGluAsnSerAlaValProValGluGlu 215
 Db ATCCCATCTGAGAACAGAGAACTCTGCGGTACCACTGAGAGAA 1020

RESULT 2
 AX039100 1261 bp DNA linear PAT 18-NOV-2000
 LOCUS Sequence 4 from Patent W0063367.
 DEFINITION AX039100
 ACCESSION AX039100
 VERSION AX039100.1 GI:11229276
 KEYWORDS
 ORGANISM Homo sapiens (human)
 SOURCE Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
 COX, P., DIXON, A., JACKSON, A. and MORGAN, K.
 A novel family of beta sub-unit proteins from a voltage-gated sodi
 um channel, nucleic acids encoding them and therapeutic or
 diagnostic uses there of
 Patent: WO 0063367-A 4 26-OCT-2000;
 WARNER-LAMBERT COMPANY (US) ; Cambridge University Technical
 Services Limited (GB)

FEATURES
 source
 1. .1261
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 1 986-122 Length: 1261
 Score: 1124.00 Matches: 215
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-09-977-579-2 (1-215) x AX039100 (1-1261)

QY 1 MetProAlaPheAsnArgLeuPheProLeuAlaSerLeuValLeuIleTyrrTPValSer 20
 Db ATGCGCTGCTTCAATGATGATTTCCCTGGCTTCTCTCGTCTTATCTACTGGGACAGT 435
 QY 21 ValCybPheProValCybValGluValProSerGluThrGluAlaValGlnGlyAsnPro 40
 Db GTCTGCTTCCCTGTGTGTGTGAAGTGCCCTCGAGAGCGAGGCGGTGACAGGACCC 495
 QY 41 MetIleLeuArgCysIleSerCysMetIleAspGluGluValGluAlaThrThrValVal 60
 Db ATGAAGCTGGCGTGCATCTCTCCATGAAAGAGAGAGAGAGGCGCACACGCGTGTG 555
 QY 61 GluTrpPheTyrrArgProGluGlyGlyIleAspPheLeuIleTyrrGluTyrrArgAnGly 80
 Db GAATGGTTCTTACAGGCCCGAGGGCGGTAAAGATTTCTTATTTACGAGTATCGGAATGGC 615
 QY 81 HisGlnGluValGluSerProPheGlnGlyArgLeuGlnTrpAsnGlySerIleAspLeu 100
 Db CACCAAGAGGTGAGAGCCCTTTCAGGGGCGCTGCGAGTGGACAGACCTG 675
 QY 101 GlnAspValSerIleThrValIleAsnValThrLeuAsnAspSerGlyLeuTyrrThrCys 120
 Db CAGAGCGTGTCCATCACTGTGCTCAACGTCACTGAGACACTCGGCTCTACACCTGC 735
 QY 121 AsnValSerArgGluPheGluPheGluAlaHisArgProPheValIleTyrrThrArgLeu 140
 Db AATGTGTCCCGGAGTTTGAAGTGTGAGCGCATCGCCCTTGTGTGAAGACAGCGCGCTG 795
 QY 141 IleProLeuArgValThrGluGluAlaGlyIleAspPheThrSerValIleSerGluIle 160
 Db ATCCCTTAAGATCACCGAGAGGCTGAGAGACTTCACTCTGTGTCTCAGAAATC 855
 QY 161 MetMetTyrrIleLeuLeuValPheLeuThrLeuTrpLeuLeuIleGluMetIleTyrrCys 180
 Db ATGATGTACATCTCTGTGCTTCTTCAACCTGTGCTGCATCGAGATGATATTTGC 915
 QY 181 TyrArgIleValSerIleAspGluGluAlaHisArgProPheValIleTyrrThrArgLeu 200
 Db TACAGAAAGGTCTCAAAAGCCGAAAGGAGGAGCCCAAGAAACCGCTGACTACCTTGGC 975
 QY 201 IleProSerGluAsnIleGluAsnSerAlaValProValGluGlu 215
 Db ATCCCATCTGAGAACAGAGAACTCTGCGGTACCACTGAGAGAA 1020

RESULT 3
 HSA243396 1261 bp mRNA linear PRI 13-MAR-2000
 LOCUS HSA243396
 DEFINITION Homo sapiens mRNA for voltage-gated sodium channel beta-3 subunit
 (scn3b gene).
 ACCESSION AJ243396
 VERSION AJ243396.2 GI:7242612
 KEYWORDS scn3b gene; voltage-gated sodium channel beta-3 subunit.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
 MORGAN, K., STEVENS, E.B., SHAW, B., COX, P., DIXON, A.K., LEE, K.,
 PINNOCK, R.D., HIGHER, J., RICHARDSON, P.J., WIZUGUCHI, K. and
 JACKSON, A.P.
 beta 3: an additional auxiliary subunit of the voltage-sensitive
 sodium channel that modulates channel gating with distinct kinetics
 Proc. Natl. Acad. Sci. U.S.A. 97 (5), 2308-2313 (2000)

JOURNAL

MEDLINE 20160948
 PUBMED 10688874
 REFERENCE 2
 AUTHORS Morgan, K.
 JOURNAL Direct Submission
 TITLE Submitted (28-JUN-1999) Morgan K., Biochemistry, University of
 Cambridge, Tennis Court Road, Cambridge, CB2 1QW, UNITED KINGDOM
 REMARK Revised by (4)
 REFERENCE 3 (bases 1 to 1261)
 AUTHORS Morgan, K.
 JOURNAL Direct Submission
 TITLE Submitted (13-MAR-2000) Morgan K., Biochemistry, University of
 Cambridge, Tennis Court Road, Cambridge, CB2 1QW, UNITED KINGDOM
 COMMENT On Mar 14, 2000 this sequence version replaced gi:7160974.
 FEATURES
 source
 1..1261
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /issue_type="brain"
 1..1261
 /gene="scn3b"
 <1..375
 /gene="scn3b"
 376..1024
 /gene="scn3b"
 /function="modulates channel gating kinetics"
 /codon_start=1
 /evidence="experimental"
 /product="voltage-gated sodium channel beta-3 subunit"
 /protein_id="CA876825.1"
 /db_xref="GI:7160975"
 /db_xref="GOA:Q9NY72"
 /translation="MPAFNRLPLASLVLIVWSVCPVCVEVPSSEATVQGNPMKLR
 CISCCKEEVATVVEWYRPEGKDFLIYENRNGHVESPRQGLQNGSKDLGD
 VSLTVNLNDSGLYTCNVSREREPAHPPVKTPTLIRPTREAGDPTSVSEI
 MNTILVFLTLMLILEMTYCRKYSKAEAAQENASDYLAIPSENKENSAPVEE"
 1025..>1261
 /gene="scn3b"
 ORIGIN
 3' UTR
 Alignment Scores:
 Pred. No.: 1.98e-122 Length: 1261
 Score: 1124.00 Matches: 215
 Percent Similarity: 100.00% Conserved: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0
 US-09-977-579-2 (1-215) x HSA243396 (1-1261)
 QY 1 MetProAlaPheAsnAArgLeuPheProLeuAlaSerLeuValLeuIleYrTrpValSer 20
 DB 376 ATGGCTGCTTCATATGATTGTTCCCTGGCTTCCTCGGCTATCTACTGCTCACT 435
 QY 21 ValCysePheProValCyseValGluValProSerGluThrGluValValGlnGlyAsnPro 40
 DB 436 GTCTGCTTCCCTGCTGTGTGTGGAAGTCCCTCGAGACGAGGCGCGTGCAGGCAACCC 495
 QY 41 MetLeuLeuArgCyseIleSerCyseMetLeuArgGluGluValGluValIleThrValVal 60
 DB 496 ATGAAGCTGGGCTGCTCTCTGATGAAGAGAGAGAGAGGAGGCGCACACCGCTGCTG 555
 QY 61 GluTrpPheYrArgProGluGluGlyLysAspPheLeuIleYrGluYrArgAsnGly 80
 DB 556 GAATGGTTCTACAGCCCGAGGCGGTAAAGATTTCCTATTATTCAGAGTATCGGAATGCG 615
 QY 81 HisGlnGluValGluSerProPheGlnGlyArgLeuGlnTrpAsnGlySerLysAspLeu 100
 DB 616 CACACGAGAGGTGAGACCCCTTTCAAGGGGCGCTGACGTGAGTGCAGACAGACACTG 675
 QY 101 GlnAspValSerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuYrThrCys 120

DB 676 CAGACGTGTCTACCTACTGCTGCTCAAGCTGCTGACACCTTGCCTCTACACCTGC 735
 QY 121 AsnValSerArgGluPheGluPheGluAlaHisArgProPheValYrThrThrArgLeu 140
 DB 736 AATGTGCTCCGGAGGTGATTGATTGAGCCCATCGGCTTTGTGGAAGACGACGGGCTG 795
 QY 141 IleProLeuArgValThrGluGluValGlyGluAspPheThrSerValValSerGluIle 160
 DB 796 ATCCCTTAAGAGTACCGAGGAGGCTGAGAGAGACTTCACTCTGTGCTTCAGAAATC 855
 QY 161 MetMetYrIleLeuLeuValPheLeuThrLeuTrpLeuLeuIleGluMetIleYrCys 180
 DB 856 ATGATGTACTCTCTTCTGCTCTTCTCACTGCTGCTGCTCATCGATGATATATTGC 915
 QY 181 TyrArgLysValSerLysAlaGluGluValAlaGlnGluAsnAlaSerAspYrLeuAla 200
 DB 916 TACGAAAGGTCTCAAAAGCCGAAAGAGGACCCCAAGAAAACGGGTCTGACTACTGCTGC 975
 QY 201 IleProSerGluLeuLysGluAsnSerAlaValProValGluGlu 215
 DB 976 ATCCCATGTGAGAACAGAGAACTCTGCGGTACCACTGAGAGAA 1020
 RESULT 4
 LOCUS AB032984 5306 bp mRNA linear PRI 11-NOV-1999
 DEFINITION Homo sapiens mRNA for KIAA1158 protein, partial cds.
 ACCESSION AB032984
 VERSION AB032984.1 GI:6330135
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 Hiroseawa, M., Nagase, T., Ishikawa, K., Kikuno, R., Nomura, N. and
 Ohara, O.
 Characterization of cDNA clones selected by the Genemark analysis
 from size-fractionated cDNA libraries from human brain
 DNA Res. 6 (5), 329-336 (1999)
 10574461
 2 (bases 1 to 5306)
 Ohara, O., Nagase, T. and Kikuno, R.
 Direct Submission
 Submitted (04-OCT-1999) Osamu Ohara, Kazusa DNA Research Institute,
 Laboratory of DNA Technology, 1532-3 Yana, Kisarazu, Chiba
 292-0812, Japan (E-mail: cdnainfo@kazusa.or.jp,
 URL: http://www.kazusa.or.jp/huge/, Tel: +81-438-52-3913,
 Fax: +81-438-52-3914)
 FEATURES
 source
 1..5306
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="h100081"
 /sex="male"
 /issue_type="brain"
 /clone_lib="PbluescriptII SK plus"
 /dev_stage="adult"
 1..5306
 /gene="KIAA1158"
 <1..695
 /gene="KIAA1158"
 /note="Start codon is not identified."
 /codon_start=3
 /product="KIAA1158 protein"
 /protein_id="BA86472.1"
 /db_xref="GI:6330136"
 /translation="RSMGSELSGSRPQKPAFNRLPLASLVLIVWSVCPVCVEV
 PSRETAVOGNPMKLR CISCCKEEVATVVEWYRPEGKDFLIYENRNGHVESPRQGLQNGSKDLGD
 VSLTVNLNDSGLYTCNVSREREPAHPPVKTPTLIRPTREAGDPTSVSEI
 MNTILVFLTLMLILEMTYCRKYSKAEAAQENASDYLAIPSENKENSAPVEE"
 TEEAGDPTSVSEIMNTILVFLTLMLILEMTYCRKYSKAEAAQENASDYLAIP

ORIGIN ENKENSAPVEE"

Alignment Scores:

| Pred. No.: | 1,26e-121 | Length: | 5306 |
|------------------------|-----------|---------------|------|
| Score: | 1124.00 | Matches: | 215 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 100.00% | Indels: | 0 |
| DB: | 9 | Gaps: | 0 |

US-09-977-579-2 (1-215) x AB032984 (1-5306)

QY 1 MetProAlAPheAsnArgLeuPheProLeuAlaSerLeuValIleuIleTyrTrpValSer 20

DB 48 ATGCTGCTCCCTTCAATGATTTGTTCCCTGCTTCTCTGCTTATCTACTGAGTCAAGT 107

QY 21 ValCysPheProValCysValGluValProSerGluThrGluAlaValGlnGlyAsnPro 40

DB 108 GTCGTCTCCCTGTGTGTGTGAAGTGCCTCGAGACGGAGCCGTGACGGGCAACCC 167

QY 41 MetLeuLeuArgCysIleSerCysMetIlyAsrGluGluValGluAlaThrThrValVal 60

DB 168 ATGAAGCTGCGCTGCATCTCTGCATGAAGAGAGAGAGGTGAGGCCACACAGGTGTG 227

QY 61 GluTrpPheTyrArgProGluGlyGlyAspPheLeuIleTyrGluTyrArgAsnGly 80

DB 228 GAATGTTCTACAGGCCCGGAGGCGGTAAAGATTCTTATTTACAGATTCGGAATGGC 287

QY 81 HisGlnGluValGluSerProPheGlnGlyArgLeuGlnTrpAsnGlySerIlyAspLeu 100

DB 288 CACACAGAGGTGAGAGACCCCTTTCAGGGGCGCTCAGATGGAATGACAGACGACCTG 347

QY 101 GlnAspValSerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuTyrTrpCys 120

DB 348 CAGGAGGTGCATCACTGTGCTCAACGTCACTTAACACCTGACCTGACCTGACCTG 407

QY 121 AsnValSerArgGluPheGluPheGluAlaHisArgProPheValIlyThrThrArgLeu 140

DB 408 AATGTTCCCGGAGGTTTGAAGTTCAGGCCCATCGGCCCTTGTGAGAGACGCGGCTG 467

QY 141 IleProLeuArgValThrGluGluAlaGlyGluAspPheThrSerValIlySerGluIle 160

DB 468 ATCCCTTAAGATCCAGGAGGAGGCTGAGAGAGACTTCACTCTGTGCTCAGAAATC 527

QY 161 MetMetTyrIleLeuLeuValPheLeuThrLeuThrLeuLeuIleGluMetIleTyrCys 180

DB 528 ATGATGTACATCTTGTGCTTCTCCTCCTGCTGCTGCTCAATCGAGATGATATATTC 587

QY 181 TTYArgIlyValSerIlyAsnIlyGluAlaAlaGlnGluAsnAlaSerAspTyrLeuAla 200

DB 588 TACAGAAAGGTCTCAAAAGCCGAAAGAGGAGGCCCAAGAAACCGCTGACTACCTTGGC 647

QY 201 IleProSerGluAsnIlyGluAsnSerAlaValProValGluGlu 215

DB 648 ATCCCATCTGAGAAACAAGAGAACTCTGGGCTACCACTGAGAGGAA 692

RESULT 5

AB097521 3296 bp mRNA linear PRI 06-DEC-2002

LOCUS AB097521

DEFINITION Macaca fascicularis brain cDNA clone:QmoA-13657, full insert sequence.

ACCESSION AB097521

VERSION AB097521.1 GI:26449236

KEYWORDS oligo capping, f54 (full insert sequence).

SOURCE Macaca fascicularis (crab-eating macaque).

ORGANISM Macaca fascicularis

REFERENCE Osada,N., Hida,M., Kusuda,J., Tanuma,R., Iseki,K., Hirata,M., Suto,Y., Hirai,M., Terao,K., Suzuki,Y., Sugano,S. and Hashimoto,K.

TITLE Assignment of 118 novel cDNAs of cynomolgus monkey brain to human chromosomes

JOURNAL Gene 275 (1), 31-37 (2001)

MEDLINE 21458551

PUBMED 11574149

REFERENCE 2 (bases 1 to 3296)

AUTHORS Hashimoto,K., Osada,N., Hida,M., Kusuda,J. and Sugano,S.

TITLE Direct Submission

JOURNAL Submitted (05-DEC-2002) Kanezumi Hashimoto, National Institute of Infectious Diseases, Division of Genetic Resources; 23-1, Toyama 1-chome, Shinjuku-ku, Tokyo 162-8640, Japan (E-mail:khashi@nih.go.jp, URL:http://www.nih.go.jp/yoken/genebank/, Tel:81-3-5285-1111(ex.2120), Fax:81-3-5285-1181)

COMMENT Lab host: TOPI0

Vector: pME18S-FL3 (Acc.No. AB009864)

R. Site1: DraIII (CACTGTTG)

R. Site2: DraIII (CACTGTTG)

Description: 1st strand cDNA was primed with an oligo(dT) primer [ATGAGCCCTTTTCTTTTCTTTT]; double-stranded cDNA was synthesized using specific 5' and 3' primers and amplified by PCR. The PCR product was digested with SfiI and size selection was performed to exclude fragments <1.5kb. The SfiI-digested PCR product was cloned into distinct DraIII sites of pME18S-FL3. XhoI sites just outside the DraIII sites can be used to isolate the cDNA insert. Libraries were constructed by oligo-capping method. Custom primers used for sequencing (5' end primer [CTTCTGCTTAAAGTGGC]; 3' end primer [CGACCTGACCTGACACAC]).

FEATURES

source

1..3296

/organism="Macaca fascicularis"

/mol_type="mRNA"

/db_xref="taxon:9541"

/clone="QmoA-13657"

/sex="male"

/tissue_type="medulla oblongata"

/clone_lib="macaque brain cDNA library QmoA"

/dev_stage="adult"

6..653

/codon_start=1

/product="hypothetical protein"

/protein_id="BAC41746.1"

/db_xref="GI:26449237"

/translation="MPAENLRPLPLVLYASVCFPVCEVSETEAVQGNPMKLR CISKMRBEVATTVVEMFPEGEGKDFILYRKHGHVESPFQGRLOMNGSKIOD VSIIVNTLNDGSLYTCNVSRFEFRAEPVKTTRLILPLRVEAGDFTSVSEI MMYIILVFTLLILMIYCYRKVKAEEAENSDYLAIRSENKENSAPVEE"

CDS

ORIGIN

Alignment Scores:

| Pred. No.: | 6e-121 | Length: | 3296 |
|------------------------|---|---------------|------|
| Score: | 1116.00 <td>Matches:</td> <td>213 </td> | Matches: | 213 |
| Percent Similarity: | 99.07% <td>Conservative:</td> <td>0 </td> | Conservative: | 0 |
| Best Local Similarity: | 99.07% <td>Mismatches:</td> <td>2 </td> | Mismatches: | 2 |
| Query Match: | 99.29% <td>Indels:</td> <td>0 </td> | Indels: | 0 |
| DB: | 9 | Gaps: | 0 |

US-09-977-579-2 (1-215) x AB097521 (1-3296)

QY 1 MetProAlAPheAsnArgLeuPheProLeuAlaSerLeuValIleuIleTyrTrpValSer 20

DB 6 ATGCTGCTCCCTTCAACGATTTGTTCCCTGCTTCTCTGCTTATCTACTGAGTCAAGT 65

QY 21 ValCysPheProValCysValGluValProSerGluThrGluAlaValGlnGlyAsnPro 40

DB 66 GTCGTCTCCCTGTGTGTGTGAAGTGCCTCGAGACGGAGCTGTGACGGGCAACCC 125

QY 41 MetLeuLeuArgCysIleSerCysMetIlyAsrGluGluValGluAlaThrThrValVal 60

DB 126 ATGAAGCTGCGCTGCATCTCTGCATGAAGAGAGAGAGGTGAGGCCACACAGGTGTG 185

QY 61 GluTrpPheTyrArgProGluGlyGlyAspPheLeuIleTyrGluTyrArgAsnGly 80

DB 186 GAATGTTCTACAGGCCCGGAGGCGGTAAAGATTCTTATTTACAGATTCGGAATGGC 245

| | | | |
|-----------|---|--|-----------------|
| QY | 61 | GIUtrPheTYrARgPProGLUGIYLYAspPheLeuIleTYrGLUtyrARgAsnGly | 80 |
| Db | 181 | GAGTGGTTCTACAGGCGCTGAGGCGCGTAAAGATTCTCTATATATGAGTACGGAAATGC | 240 |
| QY | 81 | HISgInGUaLgIuseRProPheGInGlyARgLeuGIntrPaenGlySerLYAspLeu | 100 |
| Db | 241 | CACCAAGAAAGTGAAGAGCCCTCTCCAAAGCGCGTCTGACGTGAAATGGAGCAAGACCTG | 300 |
| QY | 101 | GLInAspValSerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuTYrThrCys | 120 |
| Db | 301 | CAGACGATATCAATCACTGACTGAATTCGAAGTCACTTGTGAATGACTGGCTCTTACATGC | 360 |
| QY | 121 | AsnValSerARgGluPheGluPheGluAlaHisARgProPheValLYSThrThrARgLeu | 140 |
| Db | 361 | AATGTGTCCAGGAGGAGTTGGAATTTGAGGACACAGCGCTTTTGTGAACACACAGACTG | 420 |
| QY | 141 | IleProLeuARgValThrGluGluAlaGlyGluAspPheThrSerValIleSerGluIle | 160 |
| Db | 421 | ATACCTTTGCAGATCACTGAAGAGCGGAGAAAGACTTACCTCGTGGTCTCGGAATTC | 480 |
| QY | 161 | MetMetTYrIleLeuLeuValPheLeuThrLeuThrLeuLeuIleGluMetIleTYrCys | 180 |
| Db | 481 | ATGATGATACATCCCTCGGTCTTCTCACTTGCGCTGTTATVTGAGATATCATTTGC | 540 |
| QY | 181 | TYrARgLYAspValSerLYAspAlaGluGluAlaIleGInGUaAsnAlaSerAspTYrLeuAla | 200 |
| Db | 541 | TACGAAAGATCTCTTAAGCCGGAAGGACGACACAGAAATGCGTCTGACTACTTCTCT | 600 |
| QY | 201 | IleProSerGluAsnLYSGluAsnSerAlaValProValGluGlu | 215 |
| Db | 601 | ATCCCTTCAGAGAACAAAGAGAACTGTGTGTTACTGTGGAGGAA | 645 |
| RESUT.T 7 | | | |
| AR359849 | LOCUS | AR359849 | 2220 bp |
| | DEFINITION | Sequence 3 from patent US 6593565. | DNA |
| | ACCESSION | AR359849 | linear |
| | VERSION | AR359849.1 | PAT 17-AUG-2003 |
| | KEYWORDS | GI:33766659 | |
| | SOURCE | Unknown. | |
| | ORGANISM | Unknown. | |
| | REFERENCE | 1 (bases 1 to 2220) | |
| | AUTHORS | Healin,P. and Lynam,N.R. | |
| | TITLE | Vehicle interior rearview mirror assembly including an | |
| | JOURNAL | accessory-containing housing | |
| | FEATURES | Patent: US 6593565-A 3 15-Jul-2003; | |
| | source | Location/Qualifiers | |
| | | 1..2220 | |
| | | /organism="Unknown" | |
| | | /mol_type="genomic DNA" | |
| ORIGIN | | | |
| | Alignment Scores: | | |
| | Pred. No.: | 7.14e-120 | Length: 2220 |
| | Score: | 1105.00 | Matches: 211 |
| | Percent Similarity: | 98.14% | Conservative: 0 |
| | Best Local Similarity: | 98.14% | Mismatches: 4 |
| | Query Match: | 98.31% | Indels: 0 |
| | DB: | 6 | Gaps: 0 |
| | US-09-977-579-2 (1-215) x AR359849 (1-2220) | | |
| QY | 1 | MetProAlaPheAsnARgLeuPheProLeuAlaSerLeuValLeuIleTYrTYrValSer | 20 |
| Db | 363 | ATGCTGCTTCAACAGATTGCTTCCCTTACTCTCTAAGTCATCTACTACTGAGTCA | 422 |
| QY | 21 | ValCysPheProValCysValGluValProSerGluThrGluAlaValGInGlyAsnPro | 40 |
| Db | 423 | GTCGTCTTCCCTGTGTGTGTGTGAAGTGCCTCCGAGACACAGAGCGGTGCAGGCGAATCC | 482 |
| QY | 41 | MetLYsLeuARgCysIleSerCysMetLYsARgGluGluValGluAlaThrThrValVal | 60 |

```

Db      483 ATGAAGCTGAGTGTCATCTCTGATGAAGAGGAGAGGTGAGGCGACCATCTGTGTS 542
Qy      61 G|uTTPheTYrArpProgluglygLYvAspPheleuileTYrgIuTYrArpAsngly 80
Db      543 GAGTGTTCCTACAGGCGCTGAGGCGGTAAGATTCTTATATATAGATGCGAATGGC 602
Qy      81 H|eG|nglVAlglUserProPhegInglYArpLeuG|uTTPheAnglYserLYvAspLeu 100
Db      603 CACGAGAAAGTGAAGAGCCCTTCGAAAGCGCTGAGAGATGGAAGAGCAAGACTTG 662
Qy      101 G|uAspVAlSerIleThrValIleuAsnValThrIleuAsnAspSerGlyLeuTYrThrCys 120
Db      663 CAGAGCGTATCCATCACTGACTGACTCAATGTCACCTTGATGATGACTGCGCTTACATGC 722
Qy      121 AsnValSerArgIuPhegIuPhegIuAlaH|sArgProPheValYsThrThrArgLeu 140
Db      723 AATGTGTCCAGGAGTTCGAATTGAGAGCACAGAGGCTTTTGTGAAGACAGAGACTG 782
Qy      141 I|eProleuArgValThrgIuGlUAlaG|yglUAspPheThSerValIleSerGluIle 160
Db      783 ATACCTTTGGAGTCACTGAAGAGCGGAGAAAGACTTCCTCCGTGCTCGAATC 842
Qy      161 MetMetTYrIleleuLeuValPheleuThrLeuThrPheleuileglUmetIleTYrCys 180
Db      843 ATGATGTACATCTCTCGTCTTCTCTCACTTGAGGCTGTTATGATGATGATCTATTGC 902
Qy      181 TYrArgLYvAlSerLYsAlaG|uGlUAlaAlaG|nglUAsnAlaSerApTYrLeuAla 200
Db      903 TACAGAAAGTCTCTTAAGCGCGAAGAGGCGACAGAGAAATGCGTCTGACTACTTGGCT 962
Qy      201 I|eProSerG|uAsnLYsG|uAsnSerAlaValProValG|uGlu 215
Db      963 ATCCCTTCAGAGAACAGAGAACTCTGTGTACTCTGTGAGAGAA 1007

RESULT 8
AX039099      2220 bp      DNA      linear      PAT 18-NOV-2000
LOCUS      Sequence 3 from Patent WO0063367.
ACCESSION      AX039099
VERSION      AX039099.1 GI:11229275
KEYWORDS
SOURCE
ORGANISM      Rattus sp.
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
              Rattus.
REFERENCE
  1 Cox, P., Dixon, A., Jackson, A. and Morgan, K.
  A novel family of beta sub-unit proteins from a voltage-gated sodi
  um channel, nucleic acids encoding them and therapeutic or
  diagnostic uses there of
  Patent: WO 0063367-A 3 26-OCT-2000.
  WARNER-LAMBERT COMPANY (US) ; Cambridge University Technical
  Services Limited (GB)
FEATURES
  source
    1..2220
    /organism="Rattus sp."
    /mol_type="unassigned DNA"
    /db_xref="taxon:10118"
ORIGIN
Alignment Scores:
Pred. No.:      7.14e-120      Length:      2220
Score:          1105.00      Matches:      211
Percent Similarity: 98.14%      Conservative: 0
Best Local Similarity: 98.14%      Mismatches: 4
Query Match:    98.31%      Indels:      0
Db:             6      Gaps:      0
US-09-977-579-2 (1-215) x AX039099 (1-2220)
Qy      1 MetProAlaPheAsnArgLeuPheProLeuAlaSerLeuValleuIleTYrTPValSer 20

```

```

Db      363 ATGCTGCTCTTCAACAGATTGCTTCCCTTAGCTTCTTAGTGCTCATCTAGGCTCACA 422
Qy      21 ValCYsPheProValCYsValG|uValProSerG|uThrG|uAlaValG|nglYAsnPro 40
Db      423 GTCGTCTCCCTGTGTGTGTGAGAGTCCCTCGAAGACAGAAAGCGGTGAGGCGCAATCCC 482
Qy      41 MetLYleuArgCYsIleSerCYsMetLYsArgG|uG|uValG|uAlaThrThrValIle 60
Db      483 ATGAAGCTGAGTGTCATCTCTGATGAAGAGAGAGAGAGAGAGGCGACCATCTGTGTS 542
Qy      61 G|uTTPheTYrArpProgluglygLYvAspPheleuileTYrgIuTYrArpAsngly 80
Db      543 GAGTGTTCCTACAGGCGCTGAGGCGGTAAGATTCTTATATATAGATGCGAATGGC 602
Qy      81 H|eG|nglVAlglUserProPhegInglYArpLeuG|uTTPheAnglYserLYvAspLeu 100
Db      603 CACGAGAAAGTGAAGAGCCCTTCGAAAGCGCTGAGAGATGGAAGAGCAAGACTTG 662
Qy      101 G|uAspVAlSerIleThrValIleuAsnValThrIleuAsnAspSerGlyLeuTYrThrCys 120
Db      663 CAGAGCGTATCCATCACTGACTGACTCAATGTCACCTTGATGATGACTGCGCTTACATGC 722
Qy      121 AsnValSerArgIuPhegIuPhegIuAlaH|sArgProPheValYsThrThrArgLeu 140
Db      723 AATGTGTCCAGGAGTTCGAATTGAGAGCACAGAGGCTTTTGTGAAGACAGAGACTG 782
Qy      141 I|eProleuArgValThrgIuGlUAlaAlaG|yglUAspPheThSerValIleSerGluIle 160
Db      783 ATACCTTTGGAGTCACTGAAGAGCGGAGAAAGACTTCCTCCGTGCTCGAATC 842
Qy      161 MetMetTYrIleleuLeuValPheleuThrLeuThrPheleuileglUmetIleTYrCys 180
Db      843 ATGATGTACATCTCTCGTCTTCTCTCACTTGAGGCTGTTATGATGATGATCTATTGC 902
Qy      181 TYrArgLYvAlSerLYsAlaG|uGlUAlaAlaG|nglUAsnAlaSerApTYrLeuAla 200
Db      903 TACAGAAAGTCTCTTAAGCGCGAAGAGGCGACAGAGAAATGCGTCTGACTACTTGGCT 962
Qy      201 I|eProSerG|uAsnLYsG|uAsnSerAlaValProValG|uGlu 215
Db      963 ATCCCTTCAGAGAACAGAGAACTCTGTGTACTCTGTGAGAGAA 1007

RESULT 9
RNO243395      2220 bp      mRNA      linear      ROD 28-JAN-2003
LOCUS      Rattus norvegicus mRNA for voltage-gated sodium channel beta-3
DEFINITION
  subunit.
ACCESSION      AJ243395
VERSION      AJ243395.2 GI:7242802
KEYWORDS      scnb gene; voltage-gated sodium channel beta-3 subunit.
SOURCE      Rattus norvegicus (Norway rat)
ORGANISM      Rattus norvegicus
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
              Rattus.
REFERENCE
  1 Morgan, K., Stevens, E.B., Shaw, B., Cox, P., Dixon, A.K., Lee, K.,
  Pimock, R.D., Hynes, J., Richardson, P.J., Mizuguchi, K. and
  Jackson, A.P.
  beta 3: an additional auxiliary subunit of the voltage-sensitive
  sodium channel that modulates channel gating with distinct kinetics
  Proc. Natl. Acad. Sci. U.S.A. 97 (5), 2308-2313 (2000)
JOURNAL      MEDLINE
PUBMED      1068874
AUTHORS      Morgan, K.
TITLE      Direct Substitution
SUBMITTED      (28-JUN-1999) Morgan K., Biochemistry, University of
Cambridge, Tennis Court Road, Cambridge, CB2 1QW, UNITED KINGDOM
REVISD BY      [3]
REMARK      3 (bases 1 to 2220)
REFERENCE
  Morgan, K.

```


TITLE Direct Submission
JOURNAL Submitted (09-MAR-2000) Morgan K., Biochemistry, University of Cambridge, Tennis Court Road, Cambridge, CB2 1QW, UNITED KINGDOM
COMMENT On Mar 14, 2000 this sequence version replaced gi:7161888.
FEATURES
source
1..2220
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/issue_type="Brain"
/issue_11b="lambda ZAP"
1..2220
/gene="scn3b"
/gene="scn3b"
363..1010
/gene="scn3b"
/function="modulates channel opening kinetics"
/codon_start=1
/evidence="experimental"
/product="voltage-gated sodium channel beta-3 subunit"
/protein_id="CAB76838.1"
/db_xref="GI:7161889"
/db_xref="GOA:O9JX00"
/db_xref="UniProt/TxEmul:O9JX00"
/translation="MPAFNRDLPLASLVLIYWRVCPVCVPSSETBAVOGNPMKLR
CISCMKEEVATTVEMFPRPEGKDFLIYEYNGHQVESPFQGLQWNGSKDLQD
VSIPLAVNTLNDGLYTCNVSREPEFAHRFPVFTTRLIPLRVTEAGBPTSVSSEI
MMYLIVFLTLMLFIEMIYCYRKVSKAEAAQENASDYLAIPSENKENVVPEE"
1011..2220
/gene="scn3b"

ORIGIN
3' UTR

Alignment Scores:
Pred. No.: 7,14e-120 Length: 2220
Score: 1105.00 Matches: 211
Percent Similarity: 98.14% Conservative: 0
Best Local Similarity: 98.14% Mismatches: 4
Query Match: 98.31% Indels: 0
Gaps: 0

US-09-977-579-2 (1-215) x RNO243395 (1-2220)

QY 1 MetProAlaPheAsnAglLeuPheProLeuAlaSerLeuValLeuIleYrTTPValSer 20
DB 363 ATGCTGCTCCCTTCAACAGATTGCTTCCCTAGCTTCTAGGCTCATCTACTGAGTGA 422
QY 21 ValCySPheProValCySValGluValProSerGluThrGluAlaValGlnGlyAsnPro 40
DB 423 GTCTGCTTCCCTGTGTGTGTGGAGTCCCTCGAGACAGAGCGGTGCGAGGCAATCC 482
QY 41 MetValLeuArgCySileSerCySMetLysArgGluGluValGluAlaThrValVal 60
DB 483 ATGAAGCTGAGGTGCATCTCTGCATGAAGGAGGAGGTGAGGCGCACCACTGTGGT 542
QY 61 GluThrPheYrArgProGluGluGlyLysAspPheLeuIleYrGluYrArgAsnGly 80
DB 543 GAGTGTCTTCAACAGGCTGAGGCGGTAAAGATTCTTATATATGATATCGGAATGCG 602
QY 81 HisGlnGluValGluSerProPheGlnGlyArgLeuGlnTrpAsnGlySerLysAspLeu 100
DB 603 CACCAAGAAATGAGAGCCCTTCCAAAGCCGTCTGCAAGTGGAGAGCAAGAACTG 662
QY 101 GlnAspValSerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuYrThrCys 120
DB 663 CAGGACTATTCATCATCTGACTCAATGTCACTTGAATGACTCTGGCGCTCTACATGC 722
QY 121 AsnValSerArgGluPheGluPheGluAlaHisArgProPheValYrThrThrArgLeu 140
DB 723 AATGTGTCCAGGAGGTTCGAATTGAGGCAACAGGCGCTTTTGAAGACCAAGACTG 782
QY 141 IleProLeuArgValThrGluGluAlaGlyLysAspPheThrSerValValSerGluIle 160

DB 783 ATACCTTTGCGAGTCACTGAAGAGCGGAGAGAACTTACCTCCGTGCTCGGAATC 842
QY 161 MetMetYrIleLeuLeuValPheLeuThrLeuTrpLeuLeuIleGluMetIleYrCys 180
DB 843 ATGATGTACATCTCTGCTGCTTCTCTACCTGTGGGCTGTATATGATGATCATTC 902
QY 181 TyrArgLysValSerLysValGluGluAlaGlnGluAlaGlnGluAlaSerAspYrLeuAla 200
DB 903 TACGAAAGGCTCTTAAAGCCGAGAGGAGAGGACAGCAAGAAATGCGTGTACTTGTCT 962
QY 201 IleProSerGluAsnLysGluAsnSerAlaValProValGluGlu 215
DB 963 ATCCCTTCAGAGAACAGAGAACTCTGTGTACTGTGAGAGAA 1007

RESULT 10
AX048005
LOCUS AX048005 2632 bp DNA linear PAT 15-DEC-2000
DEFINITION Sequence 22 from Patent WO0069912.
ACCESSION AX048005
VERSION AX048005.1 GI:11876883
KEYWORDS
SOURCE
Rattus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1
AUTHORS Cuttle,R.A.
TITLE Gene encoding a sodium channel beta-3 subunit protein
JOURNAL Patent: WO 0069912-A 22 23-NOV-2000;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source
1..2632
/organism="Rattus sp."
/mol_type="unassigned DNA"
/db_xref="taxon:10118"

ORIGIN

Alignment Scores:
Pred. No.: 8.9e-120 Length: 2632
Score: 1105.00 Matches: 211
Percent Similarity: 98.14% Conservative: 0
Best Local Similarity: 98.14% Mismatches: 4
Query Match: 98.31% Indels: 0
Gaps: 0

US-09-977-579-2 (1-215) x AX048005 (1-2632)

QY 1 MetProAlaPheAsnAglLeuPheProLeuAlaSerLeuValLeuIleYrTTPValSer 20
DB 78 ATGCTGCTCCCTTCAACAGATTGCTTCCCTAGCTTCTAGGCTCATCTACTGAGTGA 137
QY 21 ValCySPheProValCySValGluValProSerGluThrGluAlaValGlnGlyAsnPro 40
DB 138 GTCTGCTTCCCTGTGTGTGTGGAGTCCCTCGAGACAGAGCGGTGCGAGGCAATCC 197
QY 41 MetValLeuArgCySileSerCySMetLysArgGluGluValGluAlaThrValVal 60
DB 198 ATGAAGCTGAGGTGCATCTCTGCATGAAGGAGGAGGTGAGGCGCACCACTGTGGT 257
QY 61 GluThrPheYrArgProGluGluGlyLysAspPheLeuIleYrGluYrArgAsnGly 80
DB 258 GAGTGTCTTCAACAGGCTGAGGCGGTAAAGATTCTTATATATGATATCGGAATGCG 317
QY 81 HisGlnGluValGluSerProPheGlnGlyArgLeuGlnTrpAsnGlySerLysAspLeu 100
DB 318 CACCAAGAAATGAGAGCCCTTCCAAAGCCGTCTGCAAGTGGAGAGCAAGAACTG 377
QY 101 GlnAspValSerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuYrThrCys 120
DB 378 CAGGACTATTCATCATCTGACTCAATGTCACTTGAATGACTCTGGCGCTCTACATGC 437
QY 121 AsnValSerArgGluPheGluPheGluAlaHisArgProPheValYrThrThrArgLeu 140

```

Db      438 AATGTCTCCAGGAGTTCGAATTCGAGGCACACAGCCCTTTGTGAAGACCAAGACTG 497
Qy      141 TLeProLeuArGValThrGluGluAlaGlyGluAspPheThSerValValSerGluIle 160
Db      498 ATACCTTTGGCATCTCAAGAGCGGAGGAGAACTTCACCTCCGGGCTCGGAATC 557
Qy      161 MetMetTyrIleLeuLeuValPheLeuThLeuTyrLeuLeuIleGluMetIleTyrCys 180
Db      558 ATATGATACATCTCTCGTCTCTCACTTGAGGCTGTTATATGATGATGATCTATGC 617
Qy      181 TyrArgValSerIysAlaGluGluAlaAlaGluGluAsnAlaSerAspTyrLeuAla 200
Db      618 TACAGAAAGGTCTTAAGGCCGAGAGGACGACACAGAAATCGTCTGACTACCTTGGT 677
Qy      201 TLeProSerGluAsnLysGluAsnSerAlaValProValGluGlu 215
Db      678 ATCCCTTCAGAGAACAGAACTGTGTGTAACCTGTGAGGAA 722

RESULT 11
AF378093      3107 bp  mRNA  linear  ROD 11-SEP-2002
LOCUS        Rattus norvegicus sodium channel beta 3 subunit (SCN3B) mRNA,
DEFINITION   complete cds.
ACCESSION   AF378093
VERSION     AF378093.1 GI:14165175
KEYWORDS
SOURCE      Rattus norvegicus (Norway rat)
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE    1 (bases 1 to 3107)
AUTHORS     Qu, Y., Curtis, R., Lawson, D., Gilbride, K., Ge, P., Distefano P.S.,
            Sileo-Santolucito, I., Catterall W.A. and Schenker, T.
TITLE       Differential modulation of sodium channel gating and persistent
            sodium currents by the beta1, beta2, and beta3 subunits
JOURNAL     Mol. Cell. Neurosci. 18 (5), 570-580 (2001)
MEDLINE     21919106
PUBMED      11922146
REFERENCE    2 (bases 1 to 3107)
AUTHORS     Curtis, R.
TITLE       Direct Submission
JOURNAL     Submitted (05-MAY-2001) Neurobiology, Millennium Pharmaceuticals
            Inc., 75 Sidney Street, Cambridge, MA 02139, USA
FEATURES
source
1..3107
/mol_type="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/tissue_type="dorsal root ganglion"
/dev_stage="adult"
1..3107
/gene="SCN3B"
/gene="SCN3B"
77..724
/gene="SCN3B"
/codon_start=1
/product="sodium channel beta 3 subunit"
/db_xref="GI:14165176"
/translaction="MPAFLRLPLASLVLIVYRVCPVCVEVSETEAVOQNPMLKLR
CISCKMEVEATVVEVFRPEGKDFLYEVRNGHOVESFQGRLOMNGSKDLOD
VSTIVNLINDSGLYTCNVSRPEFBAHRPFKTRRLIPLATTEAGDFNVSVEI
MKITLVFLTLMFLIEMIKYIKRYSKAEBAQENASDYLAIPEKNKENVVPEE"

ORIGIN
Alignment Scores:
Pred. No.:      1..1e-119      Length:      3107
Score:          1105.00      Matches:      211
Percent Similarity: 98.14%      Conservative: 0
Best Local Similarity: 98.14%      Mismatches: 0
Query Match:    98.31%      Indels:      4

```

```

DB:      10      Gaps:      0
US-09-977-579-2 (1-215) x AF378093 (1-3107)
Qy      1 MetProAlaPheAsnArgLeuPheProLeuAlaSerLeuValLeuIleTyrTrpValSer 20
Db      77 ATGCTGCTTCAACAGATTGCTTCCCTAGCTTCTCTAGTGCTCATCTACTGGGTGACAG 136
Qy      21 ValCysPheProValCysValGluValProSerGluThGluAlaValGluGluAsnPro 40
Db      137 GTCTGCTTCTCTGTGTGTGTGAAGTCCCTCGAGACAGAAAGCGGTGAGGCAATCCC 196
Qy      41 MetLeuLeuArgCysIleSerCysMetLysArgGluGluValGluAlaThrThrValVal 60
Db      197 ATGAACTGAGAGTGATCTCTCTGCAAGAGAGGAGAGAGTGGAGCCACACTGTGGT 256
Qy      61 GluTrpPheTyrArgProGluGlyIleYsAspPheLeuIleTyrGluTyrArgAsnGly 80
Db      257 GAGTGTCTTACAGAGGCTGAGGCGGTAAAGATTTCTTATATATGATGATCGGAATGCG 316
Qy      81 HisGluGluValGluSerProPheGluIleYsArgLeuGluThrAsnGlySerIysAspLeu 100
Db      317 CACCAAGAAAGTGAAGAGCCCTTCCAAAGCCGCTGCAAGATGAGAGCAAGACCTG 376
Qy      101 GlnAspValSerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuTyrThrCys 120
Db      377 CAGAGCTATCCATCATCTGTAATGATGATGATGATGATGATGATGATGATGATGATG 436
Qy      121 AsnValSerArgGluPheGluPheGluAlaHisArgProPheValIleThrThrArgLeu 140
Db      437 AATGTCTCAAGGAGTTCATTCATTCAGAGCACACAGGCCCTTTGTGTGAAGCACAGACTG 496
Qy      141 TLeProLeuArGValThrGluGluAlaGlyGluAspPheThSerValValSerGluIle 160
Db      497 ATTCCTTTGGCATCTCAAGAGCGGAGGAGAACTTCACTCCGCTGCTCGGAATC 556
Qy      161 MetMetTyrIleLeuLeuValPheLeuThrLeuTyrLeuLeuIleGluMetIleTyrCys 180
Db      557 ATATGATACATCTCTCTGCTTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 616
Qy      181 TyrArgValSerIysAlaGluGluAlaAlaGluGluAsnAlaSerAspTyrLeuAla 200
Db      617 TACAGAAAGGTCTTAAGGCCGAGAGGACGACACAGAAATCGTCTGACTACCTTGGT 676
Qy      201 TLeProSerGluAsnLysGluAsnSerAlaValProValGluGlu 215
Db      677 ATCCCTTCAGAGAACAGAACTGTGTGTAACCTGTGAGGAA 721

RESULT 12
AX047984      3108 bp  DNA  linear  PAT 15-DEC-2000
LOCUS        Sequence 1 from Patent WO0065912.
ACCESSION   AX047984
VERSION     AX047984.1 GI:11876881
KEYWORDS
SOURCE      Rattus sp.
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE    1
AUTHORS     Curtis, R.A.
TITLE       Gene encoding a sodium channel beta-3 subunit protein
JOURNAL     Patent: WO 0069912-A 1 23-NOV-2000;
            Millennium Pharmaceuticals, Inc. (US)
FEATURES
source
1..3108
/mol_type="Rattus sp."
/mol_type="unassigned DNA"
/db_xref="taxon:10118"

```

Alignment Scores:

```

Pred. No.: 1.1e-119 Length: 3108
Score: 1105.00 Matches: 211
Percent Similarity: 98.14% Conservative: 0
Best Local Similarity: 98.14% Mismatches: 4
Query Match: 98.31% Indels: 0
Gaps: 0
DB: 6
US-09-977-579-2 (1-215) x AX047984 (1-3108)

Qy 1 MetProAlaPheAsnArgLeuPheProLeuAlaSerLeuValLeuIleTyrTrpValSer 20
Db ATGCTGCTTCAACAGATTGCTCCCTAGCTTCTAGGCTCATCTACCTAGGCTCAGA 137
Qy 21 ValCysPheProValCysValGluValProSerGluThrGluAlaValGlnGlyAsnPro 40
Db GTCTGCTTCCCTGCTGTGTGTGGAAGTGCCTCGAGAGACAGACGCTCAGGCTCATCC 197
Qy 41 MetLysLeuArgCysIleSerCysMetLysArgGluGluValGluAlaThrThrValVal 60
Db ATGAGCTGAGGTGATCTCTCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 257
Qy 61 GluTrpPheTyrArgProGluGluGlyLysAspPheLeuIleTyrGluTyrArgAsnGly 80
Db GAGTGTCTTCAAGGCTGAGGCGGTAAAGATTTCTTATATATAGATATCGAGATGCG 317
Qy 81 HisGlnGluValGluSerProPheGlnGluValArgLeuGlnTyrPheGlnSerLysAspLeu 100
Db CACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 377
Qy 101 GlnAspValSerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuTyrThrCys 120
Db CAGAGCTATCATCATCTGCTTCTCATATCTTGAATGACTGTGGCTTCTACACATGC 437
Qy 121 AsnValSerArgGluPheGluPheGluValHisArgProPheValLysThrThrArgLeu 140
Db AATGATGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 497
Qy 141 IleProLeuArgValThrGluGluValAspGluAspPheThrSerValValSerGluIle 160
Db ATACCTTTGCAAGCTCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 557
Qy 161 MetMetTyrIleLeuLeuValPheLeuThrLeuTyrLeuLeuIleGlnMetIleTyrCys 180
Db ATGATGATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 617
Qy 181 TyrArgLysValSerLysAlaGluGluValAlaGlnGluAsnAlaSerAspTyrLeuAla 200
Db TACAGAAAGCTCTTAAAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 677
Qy 201 IleProSerGluLeuValGluAsnSerAlaValProValGluGlu 215
Db ATCCCTTCAAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 722
Db 678 ATCCCTTCAAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

RESULT 13
BC070899 3910 bp mRNA linear ROD 28-SEP-2004
LOCUS Rattus norvegicus sodium channel, voltage-gated, type III, beta,
DEFINITION BC070899
ACCESSION BC070899
VERSION BC070899.1 GI:47477789
KEYWORDS MGC.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 3910)
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Dicicenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,

```

```

TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
REMARK
COMMENT
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loguigliano, N.A., Peters, G.J.,
Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Huljk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Ketterman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E.,
Scherer, A., Schein, J.E., Jones, S.J. and Merrit, M.A.
Human and initial analysis of more than 15,000 full-length
nucleotide and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
2 (bases 1 to 3910)
Director MGC Project.
Direct Submission
Submitted (17-MAY-2004) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Howard Jacobs
CDNA Library Preparation: Express Genomics
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc.mgc@nih.gov
Akhter, N., Ayle, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooke, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, F., Legaapi, R.,
Maduro, Q.L., Masello, C., Masker, B., Mastian, S.D., McCloskey, D.C.,
McDowell, J., Pearson, R., Stantirpop, S., Thomas, P.J., Touchman, J.W.,
Tsurgon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAX Plate: 175 Row: 9 Column: 5
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 47575878.
location/qualifiers
1..3910
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="MGC:91409 IMAGE:7100476"
/tissue_type="lung, rat (Brown Norway)"
/clone_lib="NIH_MGC_231"
/lab_host="DH10B"
/notes="Vector: pExpress1"
1..3910
/gene="Scn3b"
/notes="synonym: Scn3"
/db_xref="LOCUSID:245956"
/db_xref="RGD:621657"
378..1025
/gene="Scn3b"
/codon_start=1
/product="sodium channel, voltage-gated, type III, beta"
/protein_id="AAH70899.1"
/db_xref="GI:47477790"
/db_xref="LOCUSID:245956"
/db_xref="RGD:621657"
/translaton="MPAFRLPLASLVLLYVWVCPVCEVPSSEAVQGNPMKLR
CISCMKREYVATVEMFRRPEGKDFLYEVNNGHGEVSPFGQSLQNNNGSDLDLD

```

VSITVNTLNDGLYTCNVSEFEPEAHRPFVKTRLLPLRVTEAGEDFTSVSEI
MNYILVFLTLMFLIEMIVCYRKVSKAEBAQENASDYLAIPSENKENSVPVEE"

ORIGIN

Alignment Scores:

| Pred. No.: | Length: | Matches: | Conservative: |
|------------------------|---------|-------------|---------------|
| 1,48e-119 | 3910 | 211 | 0 |
| Score: | 1105.00 | 98.14% | 98.14% |
| Percent Similarity: | 98.14% | Mismatches: | 4 |
| Best Local Similarity: | 98.31% | Indels: | 0 |
| Query Match: | 10 | Gaps: | 0 |

US-09-977-579-2 (1-215) x BC070899 (1-3910)

QY 1 MetProAlaPheAsnArgLeuPheProLeuAlaSerLeuValLeuIleTyrTrpValSer 20
 DB ATCCCTGCTTCAACAGATTGCTTCCCTTCAAGCTTCTTCTATGCTCATCTAGGTCAGA 437
 QY 21 ValCySPheProValCyValGluValProSerGluThrGluAlaValGlnGlyAsnPro 40
 DB GTCTGCTTCCCTGCTGCTGCTGGAAGTACCTCGAAGACAGAACGGTGCAGGTAATCCC 437
 QY 41 MetLysLeuArgCysIleSerCysMetLysArgGluGluValGluAlaThrThrValVal 60
 DB ATGAAGCTGAGTGATCTCTCTCATGAGAGAGAGAGAGGTCACCATGTGAGTG 557
 QY 61 GluTrpPheTyrArgProGluGluGlyLysAspPheLeuIleTyrGluTyrArgAsnGly 80
 DB GAGTGTGCTTCAACGCTGAGGGGCTAAAGATTCTTATATATAGTATCGAAAGGC 617
 QY 81 HisGlnGluValGluSerProPheGlnGlyArgLeuGlnTrpAsnGlySerLysAspLeu 100
 DB CACACGAGAGTGAAGAGCCCTTCCAGAGCGCTGAGTGAAGTGAAGCAAGACCTG 677
 QY 101 GlnAspValSerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuTyrCys 120
 DB CAGAGCTATACCATCTGCTTCAATGATGATGATGATGATGATGATGATGATGATG 737
 QY 121 AsnValSerArgGluPheGluPheGluAlaHisArgProPheValLysThrThrArgLeu 140
 DB AATGTGTCCAGGAGTTCGATTCGAGGACACAGCCCTTGTGAGACACAGACATG 797
 QY 141 IleProLeuArgValThrGluGluAlaGlyLysAspPheThrSerValLysSerGluIle 160
 DB ATACCTTGGGAGTCACTGGAAGGGGGGAGAGACTTCACTCCGCGTCTCGGAATC 857
 QY 161 MetMetTyrIleLeuLeuValPheLeuThrLeuTyrLeuLeuIleGluMetIleTyrCys 180
 DB ATGATGTACATCTCTCTGCTTCTTCACTTGTGCTGCTTGTGCTGCTGATGATCTATGC 917
 QY 181 TyrArgLysValSerLysAlaGluGluAlaGlnGluAlaGlnGluAlaSerAspTyrLeuAla 200
 DB TACAGAAAGTCTCTAAGCCGAGAGAGGACACAGAAATGCGTCTGCTGCTGCTGCT 977
 QY 201 IleProSerGluAsnLysGluAsnSerAlaValProValGluGlu 215
 DB ATCCCTTCAAGAACAGAGAACTGTGTGTAATCTGTGAGGAA 1022

RESULT 14

AK173115 4025 bp mRNA linear ROD 28-JUL-2004
 LOCUS AK173115 Mus musculus mRNA for mKIAA1158 protein.
 DEFINITION AK173115
 ACCESSION AK173115.1 GI:50510814
 VERSION FLI_CDNA.
 KEYWORDS Mus musculus (house mouse)
 SOURCE Mus musculus
 ORGANISM Mus musculus

REFERENCE
 AUTHORS Okazaki, N., Kikuno, R.F., Ohara, R., Inamoto, S., Koseki, H., Hirakawa, S., Suga, Y., Seino, S., Nishimura, M., Kashi, T., Hoshino, K., Kitanuma, H., Nagase, T., Ohara, O. and Koga, H.

TITLE

Prediction of the Coding Sequences of Mouse Homologues of KIAA Gene: IV. The Complete Nucleotide Sequences of 500 Mouse KIAA-Homologous cDNAs Identified by Screening of Terminal Sequences of cDNA Clones Randomly Sampled from Size-Fractionated Libraries

JOURNAL

2 (bases 1 to 4025)
 Okazaki, N., Kikuno, R.F., Nagase, T., Ohara, O. and Koga, H.

AUTHORS

Submitted (19-MAY-2004) Hisashi Koga, Kazusa DNA Research Institute, Laboratory for Genome Informatics, 2-6-7

COMMENT

The CREATR program supported by Japan science and technology corporation; cDNA full insert sequencing; Kazusa DNA Research Institute; cDNA library construction, clone selection and 5'- & 3'-end one pass sequencing.

FEATURES

Location/Qualifiers
 1..4025
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="mfj07349"
 /note="modified pBC SK+"
 1..4025
 /gene="mKIAA1158"
 <131..937
 /gene="mKIAA1158"
 /note="CDS is predicted by in silico analysis. Start codon is not identified."
 /codon_start=1
 /evidence=not experimental
 /product="mKIAA1158 protein"
 /protein_id="BAD32393.1"
 /db_xref="GI:50510815"

CDS

/translation="SGTSHVAQTACGCTPRPACBACPAVEADSDSRMTEAVPRGCE
 SSRNHLREKMPAFENRLLPLASLVLYWVAVCPVCEVSETEAVQVNSMKRCSICM
 KKEBEVATTVVMEFYRPEGKDFLIYERNHGEVSPPOGRLOMNSKDLQDVSLTV
 LVNTLNDGLYTCNVSEFEPEAHRPFVKTRLLPLRVTEAGEDFTSVSEIMYIL
 LVFLTLMFLIEMIVCYRKVSKAEBAQENASDYLAIPSENKENSVPVEE"

ORIGIN

Alignment Scores:

| Pred. No.: | Length: | Matches: | Conservative: |
|------------------------|---------|-------------|---------------|
| 1,35e-118 | 4025 | 210 | 0 |
| Score: | 1097.00 | 97.67% | 97.67% |
| Percent Similarity: | 97.67% | Mismatches: | 5 |
| Best Local Similarity: | 97.60% | Indels: | 0 |
| Query Match: | 10 | Gaps: | 0 |

US-09-977-579-2 (1-215) x AK173115 (1-4025)

QY 1 MetProAlaPheAsnArgLeuPheProLeuAlaSerLeuValLeuIleTyrTrpValSer 20
 DB ATCCCTGCTTCAACAGATTGCTTCCCTTCAAGCTTCTTCTATGCTCATCTAGGTCAGA 349
 QY 21 ValCySPheProValCyValGluValProSerGluThrGluAlaValGlnGlyAsnPro 40
 DB GTCTGCTTCCCTGCTGCTGCTGGAAGTACCTCGAAGACAGAACCGTGCAGGCAATTC 409
 QY 41 MetLysLeuArgCysIleSerCysMetLysArgGluGluValGluAlaThrThrValVal 60
 DB ATGAAGCTGAGTGCATCTCTGATGAGAGAGAGAGAGAGTGAAGCCACCATGTAGTG 469
 QY 61 GluTrpPheTyrArgProGluGluGlyLysAspPheLeuIleTyrGluTyrArgAsnGly 80
 DB GAGTGTGCTTCAACGCTGAGGGGCTAAAGATTCTTATATATAGTATCGAAATGCG 529
 QY 81 HisGlnGluValGluSerProPheGlnGlyArgLeuGlnTrpAsnGlySerLysAspLeu 100
 DB CACACGAGTGAAGAGCCCTTCCAGAGTCTGCTCATGAGATGAGAGCAAGACCTG 589
 QY 101 GlnAspValSerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuTyrCys 120


```
Qy 101 GlnAspValSerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuTyrThrCys 120
    |||||
Db 608 CAGGAGGTATCCATCACTGTTCTCAATGTCACCTGATGATGATGCTGGCCCTTAACATGT 667
    |||||
Qy 121 AsnValSerArgGluPheGluPheGluAlaHisArgProPheValIysThrThrArgLeu 140
    |||||
Db 668 AATGTGTCCAGGAGTTTGAGTTGGAAGCACACCGGCCCTTTGTGAAGACCAAGACTA 727
    |||||
Qy 141 IleProLeuArgValThrGluGluAlaGlyGluAspPheThrSerValValSerGluIle 160
    |||||
Db 728 ATACCCCTCGGAGTCACTGAGAGGCGGAGAAACTTCACCTCCGTGGTCTCGAATATC 787
    |||||
Qy 161 MetMetTyrIleLeuLeuValPheLeuThrLeuTrpLeuLeuIleGluMetIleTyrCys 180
    |||||
Db 788 ATGATGTACATCTCTCTGCTGCTTCTCTCACTTGAGCTGTTATGTAGATGATCTAATTGC 847
    |||||
Qy 181 TyrArgLysValSerLysAlaGluGluAlaAlaGluGluAsnAlaSerAspTyrLeuAla 200
    |||||
Db 848 TACAGAAAGGTCTCTAAGGCCGAGAGGCGAGCTCAGGAAATGCGTCTGACTACTTGCT 907
    |||||
Qy 201 IleProSerGluAsnLysGluAsnSerAlaValProValGluGlu 215
    |||||
Db 908 ATCCCTTCAGAGAACAGAGAACTCTGTGTACCCGTGGAGGAA 952
    |||||
```

Search completed: April 7, 2005, 16:17:51
Job time : 3047 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: April 7, 2005, 11:09:50 ; Search time 430 Seconds
(without alignments)
2959.871 Million cell updates/sec

Title: US-09-977-579-2
Sequence: 1 MPANRLPLASLVLYTWS.....SDYLAIPSEKNSAVPEE 215

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=x1p
-Q=/cgn2.1/USPTO.spool.p/US09977579/runat.06042005.163456.23582/app_query.fasta.1.391
-DB=Geneseq.16Dec04 -QMT=faasap -SUFPIX=ring -MINMATCH=0.1 -LOOPCT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blonum62 -TRANS=human40.ccd1
-LIST=45 -DOCALLIGN=200 -THR_SCORE=Pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pco -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09977579.OCGN.1.1.708 -runat.06042005.163456.23582 -NCPU=6 -ICPU=3
-NO MMAP -LARGEOUTRY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq.16Dec04:*

| | |
|-----|------------------|
| 1: | geneseqn1980s:* |
| 2: | geneseqn1990s:* |
| 3: | geneseqn2000s:* |
| 4: | geneseqn2001as:* |
| 5: | geneseqn2001bs:* |
| 6: | geneseqn2002as:* |
| 7: | geneseqn2002bs:* |
| 8: | geneseqn2003as:* |
| 9: | geneseqn2003bs:* |
| 10: | geneseqn2003cs:* |
| 11: | geneseqn2003ds:* |
| 12: | geneseqn2004as:* |
| 13: | geneseqn2004bs:* |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 1124 | 100.0 | 1261 | 3 | AAC67837 Human bet |
| 2 | 1124 | 100.0 | 1261 | 10 | ADB78651 Human ion |
| 3 | 1124 | 100.0 | 1261 | 10 | ACF57870 Human SCN |
| 4 | 1124 | 100.0 | 1510 | 4 | AAF84146 Human nov |
| 5 | 1124 | 100.0 | 4052 | 6 | ABA93727 Human sig |

| | | | | | |
|----|-------|------|------|----|---------------------|
| 6 | 1105 | 98.3 | 645 | 5 | AAC90601 Rat sodiu |
| 7 | 1105 | 98.3 | 2220 | 3 | AAC67836 Rat betac3 |
| 8 | 1105 | 98.3 | 2632 | 5 | AAC90602 Rat sodiu |
| 9 | 1105 | 98.3 | 3108 | 5 | AAC90600 Rat sodiu |
| 10 | 1024 | 91.1 | 1045 | 4 | AAK52345 Human pol |
| 11 | 1003 | 89.2 | 978 | 4 | AAH98320 Human EST |
| 12 | 1003 | 89.2 | 978 | 13 | ADSL1487 Human the |
| 13 | 927.5 | 82.5 | 1195 | 5 | AAH86764 DNA encod |
| 14 | 886.5 | 78.9 | 953 | 13 | ADSL10151 Human the |
| 15 | 742 | 66.0 | 471 | 2 | AAV86895 EST clone |
| 16 | 694 | 61.7 | 3531 | 5 | AAH86763 DNA encod |
| 17 | 477 | 42.4 | 657 | 5 | AAC90603 Rabbit so |
| 18 | 473 | 42.1 | 1335 | 10 | ACF57868 Human SCN |
| 19 | 473 | 42.1 | 1335 | 12 | ADQ19428 Human sof |
| 20 | 473 | 42.1 | 1414 | 8 | ABZ23837 Human vol |
| 21 | 472 | 42.0 | 1414 | 10 | ADB78640 Human ion |
| 22 | 472 | 42.0 | 1490 | 6 | ABK63747 Rat seque |
| 23 | 472 | 42.0 | 1490 | 10 | ADBS2855 Primary r |
| 24 | 472 | 42.0 | 1490 | 12 | ADQ09498 Rat sodiu |
| 25 | 465 | 41.4 | 1414 | 10 | ADB78641 Human ion |
| 26 | 408.5 | 36.3 | 621 | 5 | AAH86762 DNA encod |
| 27 | 262.5 | 23.4 | 850 | 4 | AAF57675 Rat sodiu |
| 28 | 262.5 | 23.4 | 850 | 4 | AAF30669 Sodium ch |
| 29 | 260.5 | 23.2 | 974 | 6 | AAD29622 Human bet |
| 30 | 234 | 20.8 | 407 | 4 | ABA08942 Human vol |
| 31 | 234 | 20.8 | 407 | 4 | AAK53329 Human pol |
| 32 | 230 | 20.5 | 855 | 10 | ADG15041 Human SBC |
| 33 | 218.5 | 19.4 | 358 | 10 | ACD98234 Human col |
| 34 | 167 | 14.9 | 1029 | 9 | ACF25382 Rat schwa |
| 35 | 165.5 | 14.7 | 3583 | 12 | ADQ23651 Human sof |
| 36 | 163.5 | 14.5 | 777 | 6 | ABA02395 Human mye |
| 37 | 163.5 | 14.5 | 1693 | 12 | ADP45442 Human vas |
| 38 | 161.5 | 14.4 | 162 | 12 | ADP66944 Human adl |
| 39 | 160.5 | 14.3 | 970 | 4 | AAF27741 Human tra |
| 40 | 160.5 | 14.3 | 983 | 4 | AAD02925 Human PRO |
| 41 | 160.5 | 14.3 | 983 | 6 | ABK33654 CDNA enco |
| 42 | 160.5 | 14.3 | 983 | 8 | ABX13481 Human DNA |
| 43 | 160.5 | 14.3 | 983 | 8 | ACA68615 Novel hum |
| 44 | 160.5 | 14.3 | 983 | 9 | ABT44344 Human PRO |
| 45 | 160.5 | 14.3 | 983 | 9 | ABT44627 Human PRO |

ALIGNMENTS

| | |
|--|--|
| RESULT 1 | |
| AAC67837 | |
| ID AAC67837 standard; cDNA; 1261 BP. | |
| XX | |
| XX AAC67837; | |
| XX AC | |
| XX 15-FEB-2001 (first entry) | |
| XX | |
| XX | |
| DE Human betac3 CDNA. | |
| XX | |
| KW Human; beta sub-unit; betac3; analgesic; anticonvulsant; | |
| KW cerebroprotective; vasotropic; cardiac; nootropic; cytostatic; | |
| KW dermatological; gene therapy; voltage-gated sodium channel; pain; | |
| KW epilepsy; stroke; ischemia; heart disease; Jacobsen Syndrome; | |
| KW familial nonchromaffin paraganglioma; phenylketonuria; | |
| KW Charcot Marie Tooth disease; ss. | |
| XX | |
| OS Homo sapiens. | |
| XX | |
| XX WO200063367-A1. | |
| XX | |
| PD 26-OCT-2000. | |
| XX | |
| PF 24-FEB-2000; 2000MO-EP001783. | |
| XX | |
| XX 15-APR-1999; 99US-0129473P. | |
| XX | |
| PA (WARN) WARNER LAMBERT CO. | |
| PA (UYCA-) UNIV CAMBRIDGE TECH SERVICES LTD. | |

XX Cox P, Dixon A, Jackson A, Morgan K;
 XX MPI; 2000-665241/64.
 DR P-PSDB; AAB36002.
 XX
 PT Novel nucleic acids encoding a beta-3 subunit from a voltage-gated sodium
 PT channel, and their corresponding polypeptides, useful for detecting and
 PT treating sodium channel-associated conditions, e.g. pain, epilepsy and
 PT stroke.
 XX
 PS Claim 10; Page 70-71; 88pp; English.
 XX
 CC The present sequence is given in the claims of a specification relating
 CC to a novel family of beta sub-unit proteins from a voltage-gated sodium
 CC channel. Human and rat beta sub-units, which have been collectively
 CC identified as beta3, have been isolated. The polynucleotides and
 CC polypeptides are useful for screening for agonists and antagonists of
 CC sodium channels. The agonists, antagonists, proteins and nucleic acids
 CC may be used diagnosing of treating diseases or conditions associated with
 CC voltage-gated sodium channels, e.g. pain, epilepsy, stroke, ischemia,
 CC heart disease, Jacobson Syndrome, Familial Nonchromaffin Paraganglioma,
 CC Phenylketonuria and Charcot Marie Tooth disease
 CC
 SQ Sequence 1261 BP; 263 A; 366 C; 372 G; 260 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 2,33e-124 Length: 1261
 Score: 1124.00 Matches: 215
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 3 Gaps: 0

US-09-977-579-2 (1-215) x AAC67837 (1-1261)

QY 1 MetProAlaPheAsnArgLeuPheProLeuAlaSerLeuValLeuIleTyrTrpValSer 20
 Db ATCCCTGCGCTTCATATGATTTGTTCCCTGCTTCTCTGCTTATCTACTGCGTCAAGT 435
 QY 21 ValCysPheProValCysValGluValProSerGluThrGluAlaValGlnGlyAsnPro 40
 Db GTCTGCTTCCCTGCTGCTGTGGAAGTCCCTCGAAGCGAGGCGCGTGCAGCAACCCC 495
 QY 41 MetLeuLeuArgCysIleSerCysMetLeuArgGluGluValGluAlaThrThrValVal 60
 Db ATGAAAGCTGGCGCTGCACTCTCTGCAAGAAAGAGAGAGGAGGCGACACAGGAGTG 555
 QY 61 GluTrpPheTyrArgProGluGlyGlyAspPheLeuIleTyrGluTyrArgAsnGly 80
 Db GAATGGTTCTACAGGCCCGAGGCGGGTAAAGATTCTTATTCAGATATCGAATGCG 615
 QY 81 HisGlnGluValGluSerProPheGlnGlyArgLeuGlnIleTrpAsnGlySerLeuAspLeu 100
 Db CACACAGAGGTGAGAGGCCCTTTCAGGGCGGCTGCAAGTGAATGCAAGCAAGCACTG 675
 QY 101 GluAspValSerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuTyrTrpCys 120
 Db CAGAGAGTGTCCATCTGCTGCTCAAGTCACTGACACACTGCTGCTTACACCTGC 735
 QY 121 AsnValSerArgIlePheGluPheGluAlaHisArgProPheValIleThrTrpArgLeu 140
 Db AATGTGTCGGGAGATTGATTTGAGGCGGCATCCGCTTGTGAAGAGAGCGGCGT 795
 QY 141 IleProLeuArgValThrGluGluAlaGlyGluAspPheThrSerValValSerGluIle 160
 Db ATCCCTTAAGATGACCCGAGGAGGCGTGAAGACTTCACTCTGCTGCTCAGAAATC 855
 QY 161 MetMetTyrIleLeuLeuValPheLeuThrLeuTrpLeuLeuIleGluMetIleTyrCys 180
 Db ATATATGATACATCTTCTGCTCTTCCACCTGCTGCTCTCATCGATGATATATTCG 915
 QY 181 TyrArgIleValSerIleValGluGluAlaGlnGluValAsnValSerAspTyrLeuAla 200

Db 916 TACAGAAAGCTCTCAAAAAGCCAGAGGAGCCCAAGAAAGAGCCGTCTACTACCTTGC 975
 QY 201 IleProSerGluAsnIleGluAsnSerAlaValProValGluGlu 215
 Db 976 ATCCCATCTGAGAACACAGAGAACTCTGCGTACCGAGTGAAGGAA 1020
 RESULT 2
 ADB78651
 ID ADB78651 standard; cDNA, 1261 BP.
 AC ADB78651;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Human ion channel subunit cDNA mutant SCN1AR exon 1 SEQ ID NO:22.
 XX
 KW ss; gene; mutant; ion channel; ion channel subunit; ICS; noctropic;
 KW neuroprotective; inotropic; antipyretic; antiarrhythmic; antimigraine;
 KW antidepressant; antiparkinsonian; neuroleptic; tranquiliser; analgesic;
 KW nephrotoxic; antidiabetic; ophthalmologically; epilepsy;
 KW ion channel dysfunction; human.
 XX
 OS Synthetic.
 OS Homo sapiens.
 PN MO2003008574-A1.
 XX
 PD 30-JAN-2003.
 XX
 PF 08-JUL-2002; 2002WC-AU000910.
 XX
 PR 18-JUL-2001; 2001AU-00006452.
 PR 05-MAR-2002; 2002AU-00000910.
 PR 13-MAY-2002; 2002AU-00002292.
 XX
 PA (BION-) BIONOMICS LTD.
 PA (WALL-) WALLACE R W.
 XX
 PI Mulley JC, Harkin LA, Dibbens LM, Phillips HA, Heron SE;
 PI Berkovic SF, Scheffer IE;
 XX
 DR MPI; 2003-239332/23.
 XX
 PT Identifying predisposition to an ion channel dysfunction, such as
 PT periodic paralysis, cardiac arrhythmias, migraine, Alzheimer's disease,
 PT schizophrenia, anxiety and depression, by detecting encoding-gene
 PT mutation events.
 XX
 PS Claim 6; SEQ ID NO 22; 106pp; English.
 XX
 CC The invention relates to a novel method for identifying a subject
 CC predisposed to a disorder associated with ion channel dysfunction. The
 CC method comprises ascertaining if at least one of the genes encoding ion
 CC channel subunits (ICS) has undergone a mutation event so that a cDNA
 CC derived from the subject has any of 134 nucleotide sequences. The method
 CC of the invention has noctropic, neuroprotective, inotropic, antipyretic,
 CC antiarrhythmic, analgesic, antidepressant, antiparkinsonian,
 CC neuroleptic, tranquiliser, analgesic, nephrotoxic, antidiabetic, and
 CC ophthalmological activity. A polynucleotide of the invention acts as an
 CC ion channel agonist, or ion channel antagonist. The methods, isolated
 CC nucleic acids, polypeptides, antibody, selective agonist, antagonist or
 CC modulator of an ion channel, cells and genetically modified non-human
 CC animal, are useful for the diagnosis and treatment of epilepsy and/or a
 CC disorder associated with ion channel dysfunction, such as hyper- or hypo-
 CC kalemic periodic paralysis, myotonia, malignant hyperthermia,
 CC myasthenia, cardiac arrhythmias, episodic ataxia, migraine, Alzheimer's
 CC disease, Parkinson's disease, schizophrenia, hyperplexia, anxiety,
 CC depression, phobic obsessive symptoms, neuropathic pain, inflammatory
 CC pain, chronic/acute pain, Bartter's syndrome, polycystic kidney disease,
 CC dent's disease, hyperinsulinaemic hypoglycaemia of infancy, cystic
 CC fibrosis, congenital stationary night blindness and total colour
 CC blindness. The present sequence represents a mutant cDNA of the

CC invention. The sequence data for this patent is not represented in the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pat_sequences.

XX Sequence 1261 BP; 263 A; 365 C; 372 G; 261 T; 0 U; 0 Other;

Alignment Scores:

| Pred. No.: | 2,33e-124 | Length: | 1261 |
|------------------------|-----------|---------------|------|
| Score: | 1124.00 | Matches: | 215 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 100.00% | Indels: | 0 |
| DB: | 10 | Gaps: | 0 |

US-09-977-579-2 (1-215) x ADB78651 (1-1261)

```

Qy 1 MePProAlaPheAenAryGleuPheProleuAlaSerleuValleuileTyrrTpValSer 20
Db 376 ATGCTGCTTCAATAGATTGTTTCCCTGCTCTCTGCTGCTTACTTACTGAGTCACT 435
Qy 21 ValCyAPheProValCyValGluValProSerGluThrGluAlaValGlnGlyAsnPro 40
Db 436 GTCTGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 495
Qy 41 MetIleuArgCyValSerCyMetIleArgGluGluValGluAlaThrThrValVal 60
Db 496 ATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 555
Qy 61 GluTrpPheTyrrArgProGluGluGlyValAspPheleuileTyrrTpValSerGly 80
Db 556 GAATGCTTCAACGCGCCGAGGCGGCTAAAGATTCTTATTTCAGATTCGAAATGCG 615
Qy 81 HisGlnGluValGluSerProPheGlnGlyArgLeuGlnTrpAsnGlySerIleAspLeu 100
Db 616 CACAGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 675
Qy 101 GluAspValSerIleThrValleuAsnValThrleuAsnAspSerGlyLeuTyrrThCys 120
Db 676 CAGAGCTGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 735
Qy 121 AsnValSerArgGluPheGluPheGluAlaHisArgProPheValIleThrThrArgLeu 140
Db 736 AATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 795
Qy 141 IleProleuArgValThrGluGluAlaGlyIleAspPheThrSerValIleSerGluIle 160
Db 796 ATCCCCCTAAGAGTCACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 855
Qy 161 MetMetIleIleleuLeuValPheleuThrleuTrpleuLeuIleGluMetIleTyrrCys 180
Db 856 ATGATGTACATCTTCTGCTGCTTCTCTCACTGCTGCTGCTGCTGCTGCTGCTGCTG 915
Qy 181 TyrArgIleValSerIleValGluGluAlaGlnGluAlaSerleuValleuileTyrrTpVal 200
Db 916 TACGAAAGGTCTTAAAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 975
Qy 201 IleProSerGluAsnIleGluAsnSerAlaValProValGluGlu 215
Db 976 ATCCATCTGAGAACAGAGAACTTGGGCTACAGTGGAGGAA 1020

```

RESULT 3

ACF57870
 ID ACF57870 standard; cDNA; 1261 BP.
 AC ACF57870;
 XX
 DT 15-JAN-2004 (first entry)
 DE Human SCN3B protein encoding cDNA.
 XX
 XX SCN1A; sodium channel type 1 alpha-subunit; anticonvulsant; analgesic;
 KM neuroprotective; anesthetic; cyostatic; cerebroprotective; cardiant;
 KM hypotensive; gene therapy; SCN3B; human; gene; ss.

```

XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT CDS 376..1023
XX FT /tag= a
XX FT /product= "SCN3B"
XX PN W02003072751-A2.
XX PD 04-SEP-2003.
XX PP 25-FEB-2003; 2003MO-US006010.
XX PR 25-FEB-2002; 2002US-0359382P.
XX PA (UVA-) UNIV VANDERBILT.
XX PI George AL, Lossin C;
XX DR WPI; 2003-712725/67.
XX DR P-PSDB; ABR3183.
XX PT Recombinantly expressed sodium channel type 1 alpha subunit, useful in
XX screening for modulators, for treating e.g. epilepsy.
XX PS Disclosure; Page 145-147; 176pp; English.
XX SC The invention relates to a recombinantly expressed and isolated human
XX SCN1A (sodium channel type 1 alpha-subunit) (I). (I), optionally
XX incorporated into a cell, is used to screen for specific modulators,
XX potentially useful as anticonvulsant, antiepileptic, neuroprotective,
XX analgesic and/or anesthetic agents, e.g. for treating severe myoclonic
XX epilepsy of infancy, stroke, cardiac arrest, hyperkalemic paralysis,
XX motor endplate diseases, hypertension, congestive heart failure and
XX muscular dystrophy also to treat cancer (SCN1A is expressed in prostatic
XX and metastatic cancer cell lines). These activities can also be provided
XX by gene therapy vectors that express (I) or the modulators. The
XX modulators, also antibodies directed against (I), are used to detect
XX sodium channel polypeptides. The present sequence represents a human
XX SCN3B protein encoding cDNA
XX SQ Sequence 1261 BP; 263 A; 366 C; 372 G; 260 T; 0 U; 0 Other;

```

Alignment Scores:

| Pred. No.: | 2,33e-124 | Length: | 1261 |
|------------------------|-----------|---------------|------|
| Score: | 1124.00 | Matches: | 215 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 100.00% | Indels: | 0 |
| DB: | 10 | Gaps: | 0 |

US-09-977-579-2 (1-215) x ACF57870 (1-1261)

```

Qy 1 MePProAlaPheAenAryGleuPheProleuAlaSerleuValleuileTyrrTpValSer 20
Db 376 ATGCTGCTTCAATAGATTGTTTCCCTGCTCTCTGCTGCTTACTTACTGAGTCACT 435
Qy 21 ValCyAPheProValCyValGluValProSerGluThrGluAlaValGlnGlyAsnPro 40
Db 436 GTCTGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 495
Qy 41 MetIleuArgCyValSerCyMetIleArgGluGluValGluAlaThrThrValVal 60
Db 496 ATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 555
Qy 61 GluTrpPheTyrrArgProGluGluGlyValAspPheleuileTyrrTpValSerGly 80
Db 556 GAATGCTTCAACGCGCCGAGGCGGCTAAAGATTCTTATTTCAGATTCGAAATGCG 615
Qy 81 HisGlnGluValGluSerProPheGlnGlyArgLeuGlnTrpAsnGlySerIleAspLeu 100
Db 616 CACAGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 675

```

QY 101 GlnAspValSerIleThrValIleuAsnValThrIleuAsnAspSerGlyLeuTyrThrCys 120
 DB 676 CAGGAGGTGTCATCACTGTGCTCAACGTCATCTGACCACTGGCCCTCAACCTTGC 735
 QY 121 AsnValSerArgIlePheGluPheGluValIleAsnArgProPheValIleThrThrArgLeu 140
 DB 736 AATGTGTCCCGGAGTTTGAGTTTGAGGGCCATCGGCCCTTTGAAAGAGACGGGGCTG 795
 QY 141 IleProLeuArgValThrGluValIleGlyIleuAsnAspPheThrSerValIleSerGluIle 160
 DB 796 ATCCCTTAAGATCACCGGAGGAGGCTGAGAGACTTCACTGTGTCTCAGAAATC 855
 QY 161 MetMetTyrIleLeuLeuValPheLeuThrLeuThrLeuLeuIleGluMetIleTyrCys 180
 DB 856 ATGATGTACATCTTCTGTGCTTCTTCTCACTGGCTGCTCATCGAGATGATATTTGC 915
 QY 181 TyrArgIleValSerIleValIleGluValIleAsnIleGlnIleuAsnIleAspTyrIleuVal 200
 DB 916 TACAGAAAGGTCTCAAAAGCCGAAAGGAGGAGCCCAAGAAACCGGTCTGACTACCTTGGC 975
 QY 201 IleProSerGluAsnIleGluAsnSerAlaValProValIleGlu 215
 DB 976 ATCCCATCTGAGAACAGAGAACTGTGCGGTACCACTGAGAGAA 1020
 RESULT 4
 ID AAF84146 standard; cDNA; 1510 BP.
 AC AAF84146;
 DT 07-SEP-2001 (first entry)
 XX Human novel sodium channel beta-1-like subunit encoding cDNA.
 XX
 KW Sodium channel; sensory neuron specific channel; beta-1-like subunit;
 KW SNS; therapeutic; pain; analgesic; ss.
 XX
 OS Homo sapiens.
 FT Key Location/Qualifiers
 FT CDS 213..860
 FT /tag= a
 FT /product= "sodium channel beta-1-like subunit"
 XX
 PN MO200144293-A2.
 PD 21-JUN-2001.
 XX
 PF 14-DEC-2000; 2000MO-GB004802.
 XX
 PR 17-DEC-1999; 99GB-00029970.
 XX
 PA (GLAXO) GLAXO GROUP LTD.
 PI Plumpston M, Powell AJ, Samsenau P;
 XX
 DR WPI; 2001-398129/42.
 DR P-PsDB; AAB85206.
 XX
 PT Novel sub-unit for voltage-gated sodium channel proteins for producing
 PT agents useful for treating pain.
 XX
 PS Claim 4; Page 29-30; 31pp; English.
 XX
 CC The invention provides a novel beta-1-like sub-unit for voltage-gated
 CC sodium ion channel polypeptide, specifically a sensory neuron specific
 CC channel (SNS) subunit. The novel beta-1-like subunit is useful for
 CC producing a therapeutic agent which is useful treating pain in a patient.
 CC The subunit can be expressed by standard recombinant methodology. The
 CC present sequence represents a human novel sodium channel beta-1-like
 CC subunit encoding cDNA
 XX

SQ Sequence 1510 BP; 350 A; 410 C; 414 G; 336 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 3.02e-124 Length: 1510
 Score: 1124.00 Matches: 215
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0
 US-09-977-579-2 (1-215) x AAF84146 (1-1510)
 QY 1 MetProAlaPheAsnArgLeuPheProLeuAlaSerLeuValIleuIleTyrTrpValSer 20
 DB 213 ATCCCTGCTTCAATGATTTGTTTCCCTGGCTTCTGTGTTATCTACGGGCACT 272
 QY 21 ValCysPheProValCysValIleValProSerGluThrGluValIleGlnIleAsnPro 40
 DB 273 GTCTGCTTCCGTGTGTGTGGAAAGTCCCTCGAGACGAGGCGTGCAGGGCAACCC 332
 QY 41 MetIleuAsnArgCysIleSerCysMetIleValIleGluValIleThrThrValVal 60
 DB 333 ATGAAGCTCGCTGATCTCTGTCATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 392
 QY 61 GluTrpPheTyrArgProGluGlyIleValAspPheLeuIleTyrGluTyrArgAsnGly 80
 DB 393 GAATGTTCTACAGGCGCCGAGGGCGGTAAAGATTCTTATTACAGATATGGATGGC 452
 QY 81 HisGlnIleValIleGluSerProPheGlnIleValIleGluIleThrAsnGlySerIleVal 100
 DB 453 CACGAGAGGTGAG 512
 QY 101 GlnAspValSerIleThrValIleuAsnValThrIleuAsnAspSerGlyLeuTyrThrCys 120
 DB 513 CAGGAGGTGTCATCACTGTGCTCAACGTCATCTGAACGACTTGGCCCTCAACCTGC 572
 QY 121 AsnValSerArgIlePheGluPheGluValIleAsnArgProPheValIleThrThrArgLeu 140
 DB 573 AATGTGTCCCGGAGTTTGAGTTTGAGGGCCATCGGCCCTTGTGAAAGAGAGCGGGCTG 632
 QY 141 IleProLeuArgValThrGluValIleGlyIleuAsnAspPheThrSerValIleSerGluIle 160
 DB 633 ATCCCTTAAGATCACCGAGAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 692
 QY 161 MetMetTyrIleLeuLeuValPheLeuThrLeuThrLeuLeuIleGluMetIleTyrCys 180
 DB 693 ATGATGTACATCTTGTGCTTCTTCTCACTGGCTGTGCTCATCGAGATGATATTTGC 752
 QY 181 TyrArgIleValSerIleValIleGluValIleAsnIleGlnIleuAsnIleAspTyrIleuVal 200
 DB 753 TACAGAAAGGTCTCAAAAGCCGAAAGGAGGAGCCCAAGAAACCGGTCTGACTACCTTGGC 812
 QY 201 IleProSerGluAsnIleGluAsnSerAlaValProValIleGlu 215
 DB 813 ATCCCATCTGAGAACAGAGAACTGTGCGGTACCACTGAGAGAA 857
 RESULT 5
 ID ABA93727 standard; cDNA; 4052 BP.
 AC ABA93727;
 DT 30-APR-2002 (first entry)
 XX
 DE Human signal transduction cDNA clone amy2_2f18.
 XX
 KW Human; foetal brain; foetal kidney; melanoma; testis; amygdala;
 KW Gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 PN MO200198454-A2.
 XX

PD 27-DEC-2001.
 XX 25-APR-2001; 2001WO-IB002050.
 PF 25-APR-2000; 2000US-0199380P.
 PR (GEHU-) GERMAN HUMAN GENOME PROJECT.
 XX
 PA
 XX Wiemann S;
 PI WPI: 2002-055860/07.
 DR P-PSDB; ABB05689.
 XX
 XX Human cDNA sequences and clones derived from human fetal brain, fetal
 PT kidney, melanoma, testis and amygdala cDNA libraries, useful in genetic
 PT screening and therapy.
 PS Claim 1; Page 174-175; 611pp; English.
 XX
 XX The present invention describes assemblies and computer readable media
 CC comprising novel human cDNA sequences and clones derived from human
 CC foetal brain, foetal kidney, melanoma, testis and amygdala cDNA
 CC libraries. ABA93702 to ABA93766 represent human cDNA sequences from the
 CC present invention which encode the proteins given in ABB05662 to
 CC ABB05729. The human cDNA sequences and clones can be used in gene
 CC therapy. The clones may be used in a variety of applications, for example
 CC they may be used in profiling assays, for providing large arrays of human
 CC genetic material for implementing large-scale screening strategies and
 CC for treating diseases via gene therapy procedures
 XX
 SQ Sequence 4052 BP; 994 A; 1033 C; 1036 G; 989 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.23e-123 Length: 4052
 Score: 1124.00 Matches: 215
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-09-977-579-2 (1-215) x ABA93727 (1-4052)
 QY 1 MetProAlaPheAsnArgLeuPheProLeuAlaSerLeuValLeuIleTyrTyrValSer 20
 DB 804 ATGCTGCTTCAATAGATGTTTCCCTGCTTCTCGCTTATCTACTGGGTGCT 863
 QY 21 ValCysPheProValCysValGluValProSerGluThrGluAlaValGlnIleAsnPro 40
 DB 864 GTCTGCTTCTGCTGTGTGTGTGGAAGTGCCTCGAGAGCGAGCGCTGACGGCAACCC 923
 QY 41 MetIleuValArgCysIleSerCysMetIleValArgGluValGluAlaThrThrValVal 60
 DB 924 ATGAAGCTGCGCTCATCTCTCGCATGAAGAGAGAGAGAGGAGGACCAACGCTGCTG 983
 QY 61 GluThrPheTyrArgProGluGluGlyIleValAspPheLeuIleTyrGluTyrArgAsnGly 80
 DB 984 GAATGTTCTACAGCGCCGAGCGCGTAAGAATTTCTTATTTACGACTATTCGGAATGCG 1043
 QY 81 HisGlnGluValGluSerProPheGlnGlyValArgLeuGlnTyrPheGlnSerIleAspLeu 100
 DB 1044 CACGAGAGGTGAGAGAGCCCTTTCAAGGGGCGCGCTGCACTGGAATGCAAGCAAGACTG 1103
 QY 101 GlnAspValSerIleThrValIleuAsnValThrLeuAsnAspSerGlyLeuTyrThrCys 120
 DB 1104 CAGACGCTGTCATCACTGCTCAACCTCACTGCAAGCACTGCGCTCTACACCTGCG 1163
 QY 121 AsnValSerArgGluPheGluPheGluValIleHisArgProPheValIleThrThrArgLeu 140
 DB 1164 AATGTTGCTCGGAGATTGAGTTGAGGCGCATCGCCCTTTTGAAACACACAGCGGGCTG 1223
 QY 141 IleProLeuArgValThrGluGluValGlnIleValAspPheThrSerValValSerGluIle 160
 DB 1224 ATCCCTTAAGAGTCAACGAGAGAGGCTGAGAGGAGCTTCACTCTGTGTCTCAGAAATC 1283

QY 161 MetMetTyrIleLeuLeuValPheLeuThrLeuTyrLeuLeuIleGluMetIleTyrCys 180
 DB 1284 ATGATGTACATCTTCTGCTGTCTTCTCTCACTTGTGGCTGCTCATCGATGATATATTGC 1343
 QY 181 TyrArgIleValSerIleValGluValAlaGlnGluAsnAlaSerAspTyrLeuAla 200
 DB 1344 TACAGAAAGTCTCAAAAGCCGAAAGGACAGCCCAAGAAAGCGCTGACTACTGCTGCC 1403
 QY 201 IleProSerGluAsnIleValAsnSerAlaValProValGluGln 215
 DB 1404 ATCCATCTGTGAGACAGAGAACTCTGCGGTACCATGAGAGAA 1448
 RESULT 6
 AAC90601
 ID AAC90601 standard; cDNA; 645 BP.
 XX
 XX AAC90601;
 XX
 DT 13-MAR-2001 (first entry)
 XX
 XX Rat sodium channel beta3 protein A1rx94h5 partial coding sequence.
 DE
 XX Rat; sodium channel beta3 protein; A1rx94h5; pain; sleep disorder;
 KW neurodegenerative disorder; mood disorder; muscle contraction; ss.
 XX
 OS Rattus sp.
 XX
 PN MO200069912-A1.
 XX
 PD 23-NOV-2000.
 XX
 PF 12-MAY-2000; 2000WO-US013144.
 XX
 PR 14-MAY-1999; 99US-0134198P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Curtis RAJ;
 DR WPI: 2001-122743/13.
 XX
 PT New rat sodium channel beta-3 subunit gene isolated from a rat dorsal
 PT root ganglion cDNA library for use in chromosome mapping, forensic
 PT medicine, monitoring clinical trials and therapeutics.
 PS Claim 1; Page 40-41; 145pp; English.
 XX
 CC The present invention provides the protein and coding sequences of the
 CC rat sodium channel beta3 protein, designated A1rx94h5. This protein is
 CC involved in the generation of pain and other sensory or perceptible nerve
 CC impulses, in the establishment and endurance of mood, neurodegenerative
 CC and sleep disorders, and in the control of muscle contraction, including
 CC movements such as the heartbeat, digestion and vascular tone. The
 CC sequences can be used in predictive medicine, screening and diagnostic
 CC assays, and in pharmacogenomics
 XX
 SQ Sequence 645 BP; 155 A; 154 C; 181 G; 155 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.7e-122 Length: 645
 Score: 1105.00 Matches: 211
 Percent Similarity: 98.14% Conservative: 0
 Best Local Similarity: 98.14% Mismatches: 4
 Query Match: 98.31% Indels: 0
 DB: 5 Gaps: 0
 US-09-977-579-2 (1-215) x AAC90601 (1-645)
 QY 1 MetProAlaPheAsnArgLeuPheProLeuAlaSerLeuValLeuIleTyrTyrValSer 20
 DB 1 ATGCTGCTTCAACAGATGCTTCCCTGCTTCTCTAGTGCATCACTACTGGGTGCA 60

```

QY 21 ValCysePheProValCyseValGluValProSerGluThrGluAlaValGlnGlyAsnPro 40
DB 61 GTCTGCTCCCTGCTGCTGTGGAAGTCCCTCCGAGACACAGCGGTGCGAGGCAATCCC 120
QY 41 MetLysLeuArgCysIleSerCysMetLysArgGluGluValGluAlaThrThrValVal 60
DB 121 ATGAGCTGAGGTGCATCTCTGTCATGAGAGGAGGAGGTGAGGCGCACACACTGTGTG 180
QY 61 GluTrpPheTyrArgProGluGluGlyLysAspPheLeuIleTyrGluTyrArgAsnGly 80
DB 181 GAGTGGTTCACAGGCTGAGGGGGGTAAGATTTCTTAATATATGATGATCGAATGCG 240
QY 81 HisGlnGluValGluSerProPheGlnGlyArgLeuGlnTrpAsnGlySerLysAspLeu 100
DB 241 CACGAGAACTGGAGAGCCCTTCACAGGCGCTGTGCAGTGGAATGGAGCAAGACCTG 300
QY 101 GlnAspValSerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuTyrThrCys 120
DB 301 CAGGACCTATCCATCACTGACTCAATGTCATCTTGATGATCACTTGCGCTCTACATGC 360
QY 121 AsnValSerArgGluPheGluPheGluAlaHisArgProPheValLysThrThrArgLeu 140
DB 361 AATGTGTCCAGGAGTTCGAATTCAGAGGCAACAAGGCTTTGTGAAGACACAGACTG 420
QY 141 IleProLeuArgValThrGluGluAlaGlyLysAspPheThrSerValValSerGluIle 160
DB 421 ATACCTTTGGCAGTCACTGAAAGGGCGGAGAAAGACTTCCTCGGTGCTCGGAAATC 480
QY 161 MetMetTyrIleLeuLeuValPheLeuThrLeuTyrLeuLeuIleGluMetIleTyrCys 180
DB 481 ATGATGATCACTCTCTGCTGCTCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 181 TyrArgLysValSerLysAlaGluGluAlaGlnGluAsnAlaSerAspTyrLeuAla 200
DB 541 TACGAAAGGTCTTAAAGCCGAAAGGCGACACAGGAAATGCGTGAATCTTGCT 600
QY 201 IleProSerGluAsnLysGluAsnSerAlaValProValGluGlu 215
DB 601 ATCCCTTCAGAGAACAGAGAACTCTGTGTACTCTGTGAGGAA 645

```

RESULT 7

```

AAC67836
ID AAC67836 standard; cDNA; 2220 BP.
XX
AC AAC67836;
XX
DT 15-FEB-2001 (first entry)
XX
DE Rat beta3 subunit cDNA.
XX
KW Rat; beta sub-unit; beta3; analgesic; anticonvulsant; cerebroprotective;
XX vasotropic; cardiant; nootropic; cytostatic; dermatological;
XX gene therapy; voltage-gated sodium channel; pain; epilepsy; stroke;
XX ischaemia; heart disease; Jacobsen Syndrome;
XX familial nonchromatin paraganglioma; phenylketonuria;
XX Charcot Marie Tooth disease; ss.
XX
OS Rattus sp.
XX
PN MO200063367-A1.
XX
PD 26-OCT-2000.
XX
PF 24-FEB-2000; 2000MO-EP001783.
XX
PR 15-APR-1999; 99US-0129473P.
XX
PA (WARN) WARNER LAMBERT CO.
XX (UYCA-) UNIV CAMBRIDGE TECH SERVICES LTD.
XX
PI Cox P, Dixon A, Jackson A, Morgan K;
XX
DR WPI, 2000-665241/64.

```

```

DR P-PSDB; AAB36001.
XX
PT Novel nucleic acids encoding a beta-3 subunit from a voltage-gated sodium
PT channel, and their corresponding polypeptides, useful for detecting and
PT treating sodium channel-associated conditions, e.g. pain, epilepsy and
PT stroke.
XX
PS Claim 6; Page 69-70; 88pp; English.
XX
CC The present sequence is given in the claims of a specification relating
CC to a novel family of beta sub-unit proteins from a voltage-gated sodium
CC channel. Human and rat beta sub-units, which have been collectively
CC identified as beta3, have been isolated. The polynucleotides and
CC polypeptides are useful for screening for agonists and antagonists of
CC sodium channels. The agonists, antagonists, proteins and nucleic acids
CC may be used diagnosing of treating diseases or conditions associated with
CC voltage-gated sodium channels, e.g. pain, epilepsy, stroke, ischaemia,
CC heart disease, Jacobsen Syndrome, familial Nonchromatin Paraganglioma,
CC Phenylketonuria and Charcot Marie Tooth disease
XX
SQ Sequence 2220 BP; 573 A; 557 C; 561 G; 529 T; 0 U; 0 Other.
XX
Alignment Scores:
Pred. No.: 9,88e-122 Length: 2220
Score: 1105.00 Matches: 211
Percent Similarity: 98.14% Conservative: 0
Best local Similarity: 98.14% Mismatches: 4
Query Match: 98.31% Indels: 0
DB: 3 Gaps: 0
XX
US-09-977-579-2 (1-215) x AAC67836 (1-2220)
QY 1 MetProAlaPheAsnArgLeuPheProLeuAlaSerLeuValLeuIleTyrTrpValSer 20
DB 363 ATGCTGCTCCCTTCAACGATTTGCTTCCCTGACTTCTTAAGTCTCACTAGTGGTCAAG 422
QY 21 ValCysePheProValCyseValGluValProSerGluThrGluAlaValGlnGlyAsnPro 40
DB 423 GTCTGCTCCCTGCTGCTGTGGAAGTCCCTCGGAGACAGAGCGGTGCGAGGCAATCCC 482
QY 41 MetLysLeuArgCysIleSerCysMetLysArgGluGluValGluAlaThrThrValVal 60
DB 483 ATGAACTGAGGTGATCTCTGTCATGAAAGAGGAGAGAGGTGAGGCCACCACTGTGGTG 542
QY 61 GluTrpPheTyrArgProGluGluGlyLysAspPheLeuIleTyrGluTyrArgAsnGly 80
DB 543 GAGTGGTTCACAGGCTGAGGGCGGTAAAGATTTCTTAATATATGATGATCGAATGCG 602
QY 81 HisGlnGluValGluSerProPheGlnGlyArgLeuGlnTrpAsnGlySerLysAspLeu 100
DB 603 CACCAAGAAAGTGGAGAGCCCTTCACAGGCGCGTCTGCAGTGGAATGGAGCAAGACCTG 662
QY 101 GlnAspValSerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuTyrThrCys 120
DB 663 CAGGAGGTATCCATCACTGACTCAATGTCATCTTGATTAATGACTCTGCGCTTACATGC 722
QY 121 AsnValSerArgGluPheGluPheGluAlaHisArgProPheValLysThrThrArgLeu 140
DB 723 AATGTGTCCAGGAGTTCGAATTCGAGGCAACAGGCTTTGTGAAGCACAGACTG 782
QY 141 IleProLeuArgValThrGluGluAlaGlyLysAspPheThrSerValValSerGluIle 160
DB 783 ATACCTTTGGCAGTCACTGAAAGGGCGGAGAAAGACTTCACCTCGGTGCTCGGAAATC 842
QY 161 MetMetTyrIleLeuLeuValPheLeuThrLeuTyrLeuLeuIleGluMetIleTyrCys 180
DB 843 ATGATGATCACTCTCTGCTGCTTCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 902
QY 181 TyrArgLysValSerLysAlaGluGluAlaGlnGluAsnAlaSerAspTyrLeuAla 200
DB 903 TACGAAAGGTCTTAAAGCCGAAAGGCGACACAGGAAATGCGTGTACTACTCTTGCT 962
QY 201 IleProSerGluAsnLysGluAsnSerAlaValProValGluGlu 215

```

Db 963 ATCCCTTCAGAGAACAGAGAACTCTGTGCTGCTGAGAGAA 1007

RESULT 8
AAC90602
ID AAC90602 standard; DNA; 2632 BP.
XX AAC90602;
AC AAC90602;
DT 13-MAR-2001 (first entry)
XX
XX Rat sodium channel beta3 protein Alrx94h5 related sequence.
DE
XX Rat; sodium channel beta3 protein; Alrx94h5; pain; sleep disorder;
KM neurodegenerative disorder; mood disorder; muscle contraction; ds.
XX
XX Rattus sp.
OS
PN WO200069912-A1.
XX
XX PD 23-NOV-2000.
XX
XX PF 12-MAY-2000; 2000WO-US013144.
XX
XX PR 14-MAY-1999; 99US-0134198P.
XX
XX PA (MILL-) MILLENNIUM PHARM INC.
XX
XX PI Curtis RAJ;
XX
XX WPI; 2001-122743/13.
XX
XX PT New rat sodium channel beta-3 subunit gene isolated from a rat dorsal
PT root ganglion cDNA library for use in chromosome mapping, forensic
PT medicine, monitoring clinical trials and therapeutics.
XX
XX PS Claim 1; Page 41-42; 145pp; English.
XX
XX CC The present invention provides the protein and coding sequences of the
CC rat sodium channel beta3 protein, designated Alrx94h5. This protein is
CC involved in the generation of pain and other sensory or perceptive nerve
CC impulses, in the establishment and endurance of mood, neurodegenerative
CC and sleep disorders, and in the control of muscle contraction, including
CC movements such as the heartbeat, digestion and vascular tone. The
CC sequences can be used in predictive medicine, screening and diagnostic
CC assays, and in pharmacogenomics
XX
XX SQ Sequence 2632 BP; 693 A; 636 C; 611 G; 692 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,266-121 Length: 2632
Score: 1105.00 Matches: 211
Percent Similarity: 98.14% Conservative: 0
Best Local Similarity: 98.14% Mismatches: 4
Query Match: 98.31% Indels: 0
DB: 5 Gaps: 0

US-09-977-579-2 (1-215) x AAC90602 (1-2632)

Qy 1 MetProAlaPheAsnArgLeuPheProLeuAlaSerLeuValLeuIleTyrTrpValSer 20
Db 78 ATGCTGCTTCACAGATGCTTCCCTAGCTTCTAGAGCTCATCTACGCGTCA 137
Qy 21 ValCysPheProValCysValGluValProSerGluThrGluAlaValGlnGlyAsnPro 40
Db 138 GTCCTCTCCCTGCTGTGTGTGGAAGTGCCTCGAGACAGAACGCGTCAGGCAATGCC 197
Qy 41 MetValLeuArgCysIleSerCysMetLysArgGluGluValGluAlaThrThraVal 60
Db 198 ATGAAGCTGAGGTGATCTCCGCAATGAGAGAGGAGGAGTGGAGGCCACCATGTGTG 257
Qy 61 GluTrpPheTyrArgProGluGluGlyLysAspPheLeuIleTyrGluTyrArgAsnGly 80

Db 258 GAGTGTCTTCAAGGCGCTGAGGCGGTAAAGATTTCCTATATATAGATATCGAATGCG 317
Qy 81 HisGlnGluValGluSerProPheGlnGlyArgLeuGlnTyrPasnGlySerLysAspLeu 100
Db 318 CACGAGAAAGTGAAGAGCCCTTCAGAGCCGCTCGAGTGAATGGAGCAAGACCTG 377
Qy 101 GluAspValSerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuTyrThrCys 120
Db 378 CAGACGATCATCATCTAGTCAATGTCATTGAAATGACTGTGGCTCTACACATGC 437
Qy 121 AsnValSerArgGluPheGluPheGluAlaHisArgProPheValLysThrThrArgLeu 140
Db 438 AATGTGTCCAGAGGATTGATTCAGAGCAACAGGCGCTTTGTGAAGACCAAGACTG 497
Qy 141 IleProLeuArgValThrGluGluAlaGlyGluAspPheThrSerValValSerGluIle 160
Db 498 ATACCTTTGCGAGTCACTAGGAAGCGGAGAACACTTCCTCGGTGCTCCGGAATC 557
Qy 161 MetMetTyrIleLeuLeuValPheLeuThrLeuThrLeuLeuIleGluMetIleTyrCys 180
Db 558 ATGATGATCATCTCTGCTGCTTCTCCTCCTTGAGCTGTTTATGATGATATCTATTGC 617
Qy 181 TyrArgLysValSerLysAlaGluGluAlaValGlnGluAsnAlaSerAspTyrLeuAla 200
Db 618 TACAGAAAGTCTCTAAGGCGGAGGAGGACACAGGAAATGGGTGACTACCTTCT 677
Qy 201 IleProSerGluAsnLysGluAsnSerAlaValProValGluGlu 215
Db 678 ATCCCTTCAGAGAACAGAGAACTCTGTGCTGCTGAGAGAA 722

RESULT 9
AAC90600
ID AAC90600 standard; cDNA; 3108 BP.
XX
XX AC AAC90600;
XX
XX DT 13-MAR-2001 (first entry)
XX
XX DE Rat sodium channel beta3 protein Alrx94h5 coding sequence.
XX
XX DE Rat; sodium channel beta3 protein; Alrx94h5; pain; sleep disorder;
KM neurodegenerative disorder; mood disorder; muscle contraction; ss.
XX
XX OS Rattus sp.
XX
XX PN WO200069912-A1.
XX
XX PD 23-NOV-2000.
XX
XX PF 12-MAY-2000; 2000WO-US013144.
XX
XX PR 14-MAY-1999; 99US-0134198P.
XX
XX PA (MILL-) MILLENNIUM PHARM INC.
XX
XX PI Curtis RAJ;
XX
XX WPI; 2001-122743/13.
XX
XX DR P-PSDB; AAB50243, AAB50245.
XX
XX PT New rat sodium channel beta-3 subunit gene isolated from a rat dorsal
PT root ganglion cDNA library for use in chromosome mapping, forensic
PT medicine, monitoring clinical trials and therapeutics.
XX
XX PS Claim 1; Fig 1; 145pp; English.
XX
XX CC The present invention provides the protein and coding sequences of the
XX rat sodium channel beta3 protein, designated Alrx94h5. This protein is
XX involved in the generation of pain and other sensory or perceptive nerve
XX impulses, in the establishment and endurance of mood, neurodegenerative
XX and sleep disorders, and in the control of muscle contraction, including
XX movements such as the heartbeat, digestion and vascular tone. The
XX sequences can be used in predictive medicine, screening and diagnostic

CC assays, and in pharmacogenomics
 XX Sequence 3108 BP; 830 A; 748 C; 727 G; 803 T; 0 U; 0 Other;

Alignment Scores:

| Pred. No.: | 1,6e-121 | Length: | 3108 |
|------------------------|----------|---------------|------|
| Score: | 1105.00 | Matches: | 211 |
| Percent Similarity: | 98.14% | Conservative: | 0 |
| Best Local Similarity: | 98.14% | Mismatches: | 4 |
| Query Match: | 98.31% | Indels: | 0 |
| DB: | 5 | Gaps: | 0 |

US-09-977-579-2 (1-215) x AAC90600 (1-3108)

```

Qy 1 MetProAlaPheAsnArgLeuPheProLeuAlaSerLeuValIleuIleTyrTrpValSer 20
Db 78 ATGCTGCTTCAACACATGCTTCCCTACGCTTCTAGGCTCATCTGAGTCCAGCA 137
Qy 21 ValCysPheProValCysValGluValProSerGluThrGluAlaValGlnGlyAsnPro 40
Db 138 GTCTGCTTCCCTGCTGTGTGTGAAGTCCCTCGAGACAGAGGCGTGCAGGGCAATCCC 197
Qy 41 MetLeuLeuArgCysIleSerCysMetLeuArgGluGluValGluAlaThrThrValVal 60
Db 198 ATGAAGCTGAGGTGCATCTCTGCTATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 257
Qy 61 GluTrpPheTyrArgProGluGluGlyValAspPheLeuIleTyrGluTyrArgAsnGly 80
Db 258 GAGTGTCTTCAACGAGGCTGAGGCGGTAAAGATTTCTTATATATGATGATCGGAATGGC 317
Qy 81 HisGlnGluValGluSerProPheGlnGlyArgLeuGlnTrpAsnGlySerLeu 100
Db 318 CACGAGAGATGAGAGAGCCCTTCCAGAGCGCTCTGACGAGGATGAGAGCAAGACCTG 377
Qy 101 GlnAspValSerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuTyrTrpCys 120
Db 378 CAGACCTATCCATCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 437
Qy 121 AsnValSerArgGluPheGluPheGluAlaHisArgProPheValIleTyrThrArgLeu 140
Db 438 AATGTGTCGAGGAGGATTCGATTCGAGGACGACGAGGCTTTGTGAGACACGAGACTG 497
Qy 141 IleProLeuArgValThrGluGluAlaGlyValAspPheThrSerValIleSerGluIle 160
Db 498 ATACCTTTGCGAGTCACTCAAGAGGCGGAGAGACTTCACTCCGCTCGGAAATC 557
Qy 161 MetMetIleIleLeuLeuValPheLeuThrLeuTrpLeuLeuIleGluMetIleTyrCys 180
Db 558 ATGATGTACATCTCTCTGCTGCTTCTTCACTTGTGCTGCTTGTATGATGATCTATTCG 617
Qy 181 TyrArgValSerIleValGluGluAlaGlnGluAsnAlaSerAspTyrLeuAla 200
Db 618 TACAGAAAGTCTCTAAGGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 677
Qy 201 IleProSerGluAsnLeuGluAsnSerAlaValProValGluGlu 215
Db 678 ATCCCTTCAAGAGCAAGAGAACTGTGTGCTGCTGAGGAA 722

RESULT 10
AAK52345
ID AAK52345 standard; cDNA; 1045 BP.
XX
AC AAK52345;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 890.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation; ss.
XX

```

OS Homo sapiens.

PN W0200157190-A2.

PD 09-AUG-2001.

PF 05-FEB-2001; 2001WO-US004098.

PR 03-FEB-2000; 2000US-00496914.

PR 27-APR-2000; 2000US-00560875.

PR 20-JUN-2000; 2000US-00598075.

PR 19-JUL-2000; 2000US-00620325.

PR 01-SEP-2000; 2000US-00654936.

PR 15-SEP-2000; 2000US-00663561.

PR 20-OCT-2000; 2000US-00693325.

PR 30-NOV-2000; 2000US-00728422.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;

PI Ma Y, Zhao Q, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;

PI Xue HJ, Yang Y, Wejrtman T, Goodrich R;

PI WPI; 2001-476283/51.

DR P-PSDB; AAM79212.

XX

PT Nucleic acids encoding polypeptides with cytokine-like activities, useful

PT in diagnosis and gene therapy.

PS Claim 1; Page 2934-2935; 6221pp; English.

XX The invention relates to polynucleotides (AAK51456-AAK53435) and the

XX encoded polypeptides (AAM78923-AAM80302) that exhibit activity elating to

XX cytokine, cell proliferation or cell differentiation or which may induce

XX production of other cytokines in other cell populations. The

XX polynucleotides and polypeptides are useful in gene therapy, vaccines or

XX peptide therapy. The polypeptides have various cytokine-like activities,

XX e.g. stem cell growth factor activity, haematopoiesis regulating

XX activity, tissue growth factor activity, immunomodulatory activity and

XX activin/inhibin activity and may be useful in the diagnosis and/or

XX treatment of cancer, leukaemia, nervous system disorders, arthritis and

XX inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111

XX (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the

XX sequence listing were missing at the time of publication

SO Sequence 1045 BP; 222 A; 289 C; 314 G; 220 T; 0 U; 0 Other;

Alignment Scores:

| Pred. No.: | 1,72e-112 | Length: | 1045 |
|------------------------|--|---------------|------|
| Score: | 1024.00 <td>Matches:</td> <td>195 </td> | Matches: | 195 |
| Percent Similarity: | 100.00% <td>Conservative:</td> <td>0 </td> | Conservative: | 0 |
| Best Local Similarity: | 100.00% <td>Mismatches:</td> <td>0 </td> | Mismatches: | 0 |
| Query Match: | 91.10% <td>Indels:</td> <td>0 </td> | Indels: | 0 |
| DB: | 4 | Gaps: | 0 |

US-09-977-579-2 (1-215) x AAK52345 (1-1045)

```

Qy 1 MetProAlaPheAsnArgLeuPheProLeuAlaSerLeuValIleuIleTyrTrpValSer 20
Db 302 ATGCTGCTTCAACATGATTTGTTCCCTGCGCTTCTCTGCTTATCTACGAGGCTCACT 361
Qy 21 ValCysPheProValCysValGluValProSerGluThrGluAlaValGlnGlyAsnPro 40
Db 362 GTCTGCTTCCCTGCTGTGTGTGAAGTCCCTCGAGACGAGGCGCTGACGAGCAATCCC 421
Qy 41 MetLeuLeuArgCysIleSerCysMetLeuArgGluGluValGluAlaThrThrValVal 60
Db 422 ATGAAGCTGCGCTGCAATCTCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 481
Qy 61 GluTrpPheTyrArgProGluGluGlyValAspPheLeuIleTyrGluTyrArgAsnGly 80
Db 482 GAGTGTCTTCAACGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 541

```

QY 81 HieglngluValIGUserProphengluValArgleugInttpaanglySerlyAspLeu 100
 DB 542 CACGAGAGGTGAGAGCCCTTTCAGGGGGCCCTGACGTGAATGGACAGACCTG 601
 QY 101 GlnAspValSerIleThrValIleuAsnValThrIleuAsnAspSerGlyLeuThrCys 120
 DB 602 CAGACGCTGTCACACCTGCTCAACCTCACTGACCTGACCTGACCTGACCTGACCTG 661
 QY 121 AsnValSerArgGluPheGluPheGluValHisArgProPheValIleThrArgLeu 140
 DB 662 AATGTGTCCCGGAGTTGAGTTGAGGCGCATCGGCCCTTGTGAAGACGACGGGCTG 721
 QY 141 IleProLeuArgValIleThrGluGluValGlyGluAspPheThrSerValIleSerGluIle 160
 DB 722 ATCCCTTAAGAGTACCGAGAGGCTGAGAGACCTGCTGCTGCTGCTGCTGCTGCTG 781
 QY 161 MetMetTyrlleLeuLeuValPheLeuThrLeuThrLeuLeuIleGluMetIleThrCys 180
 DB 782 ATGATGTACATCTCTTGTGCTCTTCTCTCACTTGTGCTGCTCATCGAGATATATTTGC 841
 QY 181 TyrArgLyValSerLyValGluGluValIleGluValIleGluValIleGluValIle 195
 DB 842 TACAGAAAGGTCTCAAAAGCCGAAAGAGGACCCCAAGAAACGCG 886
 RESULT 11
 ID AAH98320 standard; cDNA; 978 BP.
 XX AAH98320;
 AC 12-OCT-2001 (first entry)
 DT 12-OCT-2001 (first entry)
 DE Human EST-derived coding sequence SEQ ID NO: 177.
 XX Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
 KM tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
 KM diagnostics; forensic test; gene mapping; genetic disorder; biodiversity;
 KM gene therapy; nutrition; ss.
 XX Homo sapiens.
 OS
 XX WC0200154477-A2.
 PN 02-AUG-2001.
 PD 02-AUG-2001.
 XX 25-JAN-2001; 2001WO-US002687.
 PF 25-JAN-2001; 2000US-00491404.
 PR 17-JUL-2000; 2000US-00617746.
 PR 03-AUG-2000; 2000US-00631451.
 PR 15-SEP-2000; 2000US-0063870.
 XX (HYSE-) HYSEQ INC.
 PA
 XX Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Xuendi V;
 PI Cao Y, Drmanac RA, Zhang J, Weirman T;
 DR WPI; 2001-476164/51.
 DR P-PSDB; AAM23661.
 XX Isolated polypeptide for treatment of diseases, diagnostics, raising
 PT antibodies and research use.
 PT
 PS Claim 1; Page 299-300; 1275pp; English.
 XX
 CC The present invention provides the protein and coding sequences of novel
 CC proteins from a variety of organisms, including human, dog, cat, horse,
 CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
 CC urchin and tomato. These were derived from expressed sequence tags (ESTs)
 CC from the organism of interest. They can be used in diagnostics,
 CC forensics, gene mapping, identification of mutations, to assess
 CC biodiversity and for nutritional purposes. The present sequence is a cDNA
 CC of the invention

XX SQ Sequence 978 BP; 193 A; 283 C; 293 G; 209 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 5,14e-110 Length: 978
 Score: 1003.00 Matches: 196
 Percent Similarity: 93.46% Conservative: 4
 Best Local Similarity: 91.59% Mismatch: 12
 Query Match: 89.23% Indels: 2
 DB: 4 Gaps: 0
 US-09-977-579-2 (1-215) x AAH98320 (1-978)
 QY 1 MetProAlaPheAsnArgLeuPheProLeuAlaSerIleValIleuIleTyrlPValSer 20
 DB 295 ATGCTTCTCTCAATATAGATTGTTCCCTGCTCTCTCTGCTTACTCTGCTGCTAGT 354
 QY 21 ValCysPheProValCysValGluValProSerGluThrGluValIleGluValIle 40
 DB 355 GTCTGCTTCCCTGTGTGTGTGAAGTCCCTCGAGAGGAGCCGTGCAAGGCAACCC 414
 QY 41 MetLySerLeuArgCysIleSerCysMetLyArgGluGluValGluValIleValIle 60
 DB 415 ATGAGAGCTGGGCTGCTCATCTCTGATGAGAGAGAGAGGTGAGGCAACGCTGCTG 474
 QY 61 GluTrpPheTyrlArgProGluGlyGlyLyAspPheLeuIleTyrlGlyTyrlArgAsnGly 80
 DB 475 GAATGATTCTTACAGGCCCGAGGGCGGTAAAGATTCTTATTAAGGTATCGGAATGCG 534
 QY 81 HieglngluValIGUserProphengluValArgleugInttpaanglySerlyAspLeu 100
 DB 535 CACGAGAGGTGAGAGCCCTTTCAGGGGGCCCTGACGTGAATGGACAGACCTG 594
 QY 101 GlnAspValSerIleThrValIleuAsnValThrIleuAsnAspSerGlyLeuThrCys 120
 DB 595 CAGACGCTGTCACACCTGCTCAACCTCACTGACCTGACCTGACCTGACCTGACCTG 654
 QY 121 AsnValSerArgGluPheGluPheGluValHisArgProPheValIleThrArgLeu 140
 DB 655 AATGTGTCCCGGAGTTGAGTTGAGGCGCATCGGCCCTTGTGAAGAGACCGGCGCTG 714
 QY 141 IleProLeuArgValIleThrGluGluValGlyGluAspPheThrSerValIleSerGluIle 160
 DB 715 ATCCCTTAAGAGTACCGAGAGGCTGAGAGACCTGACCTTGTGCTCAAGAAATC 774
 QY 161 MetMetTyrlleLeuLeuValPheLeuThrLeuThrLeuLeuIleGluMetIleThrCys 180
 DB 775 ATGATGTACATCTCTTGTGCTCTTCTCACTTGTGCTGCTCATCGAGATGATCATTTGC 834
 QY 181 -TyrlArgLyValSerLyValGluGluValIleGluValIleGluValIleSerAspTyrlLeu-A 200
 DB 835 CTACAGAACCGGTATCAACAGACGAGCCGCCCAACAGAAACCGAGTGGCTACCTTTG 894
 QY 200 laileProserGluAsnLyGluAsnSerAlaValPro 212
 DB 895 CGATTCCATTGAGACAGGAAAATTTCTGGTACCT 932
 RESULT 12
 ID ADS11487 standard; DNA; 978 BP.
 XX ADS11487;
 AC ADS11487;
 DT 16-DEC-2004 (first entry)
 DT
 XX 16-DEC-2004 (first entry)
 DE Human therapeutic config DNA - SEQ ID 1724.
 XX
 XX antiinflammatory; neuroprotective; antianemic; cytostatic; vulnerary;
 KM inflammatory; haemotopoleis; immunity; neurodegenerative; stem cell;
 KM aplastic anaemia; cancer; wound healing; gene therapy; ds; gene.
 XX
 OS Homo sapiens.

PN WO2004080148-A2.
 XX
 PD 23-SEP-2004.
 XX
 PF 30-SEP-2003; 2003WO-US030720.
 XX
 PR 02-OCT-2002; 2002US-0416186P.
 XX
 PA (NUVE-) NUVELO INC.
 XX
 PI Tang YT, Asundi V, Ren F, Zhang J, Wehrman T, Wang Z, Ma Y,
 PI Wang D, Chen R, Zhao Qa, Wang J, Ghosh M, Xue AJ, Weng G, Zhou P,
 PI
 DR WPI; 2004-668857/55.
 DR P-PSDB; ADS12085.
 XX
 XX New polynucleotide, useful in preparing a composition for diagnosing or
 PT treating inflammatory, neurodegenerative or stem cell disorders, e.g.,
 PT aplastic anemia or cancer for promoting wound healing.
 XX
 PS Example 2; SEQ ID NO 1724; 718bp; English.
 CC The invention relates to a novel isolated polynucleotide and the encoded
 CC polypeptide. The molecules of the invention demonstrate antiinflammatory,
 CC neuroprotective, anti-naemic, cytoskeletal and vulnereary activities and may
 CC be useful in preparing a composition for diagnosing or treating
 CC inflammatory, haematopoietic, immune, neurodegenerative or stem cell
 CC disorders, such as aplastic anaemia or cancer, as well as for promoting
 CC wound healing. The molecules may also be utilised during gene therapy
 CC procedures. The current sequence is that of a human therapeutic contig
 CC DNA of the invention. The current sequence is not shown explicitly within
 CC the specification but can be accessed from the WIPO web-site.
 XX
 SO Sequence 978 BP; 193 A; 283 C; 293 G; 209 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 5.14e-110 Length: 978
 Score: 1003.00 Matches: 196
 Percent Similarity: 93.46% Conservative: 4
 Best Local Similarity: 91.59% Mismatches: 12
 Query Match: 89.23% Indels: 2
 DB: Gaps: 0
 US-09-977-579-2 (1-215) x ADS11487 (1-978)
 QY 1 MetProAlaPheAsnArgLeuPheProLeuAlaSerLeuValLeuIleTyrTTPValSer 20
 DB 295 ATGCTGCTGCTTCAATAGATTGTTCCCTGCTTCTCTGCTTATCTAGTGGTCAAGT 354
 QY 21 ValCyPheProValCyValAlGluValProSerGluThrGluAlaValGlnGlyAsnPro 40
 DB 355 GTCTGCTTCCCTGTGTGTGTGGAAGTCCCTCGAAGACGAGCCGCTGACGGCAACCCC 414
 QY 41 MetIleLeuArgCysIleSerCysMetIleValArgGluGluValAlaThrThrValVal 60
 DB 415 ATGAACTGGCGCTGCAATCTCTCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 474
 QY 61 GluThrPheTyrArgProGluGluGlyValAspPheLeuIleTyrGluTyrArgAsnGly 80
 DB 475 GAATGGTCTTACAGGGCCGAGGGCGGGAAGATTCTTATTTACAGATTCGGAATGGC 534
 QY 81 HisGlnGluValGluSerProPheGlnGlyArgLeuGlnIleThrPheAsnGlySerIleVal 100
 DB 535 CACCAAGAGGTGAGAGAGCCCTTTCAGGGGGCGCTCGAAGTGAATGAGCAAGAGACCTG 594
 QY 101 GlnAspValSerIleThrValIleAsnValThrLeuAsnAspSerGlyLeuTyrThrCys 120
 DB 595 CAGGAGGTGCTTCACTGCTCAACGTCACCTCTGACCACTCTGCTCTTCACTGCTG 654
 QY 121 AsnValSerArgGluPheGluPheGluAlaHisArgProPheValIleValThrArgLeu 140
 DB 655 AATGTGTCCTGGGAGTTGAGTTGAGGGCGCAATGGCCCTTGTGAAAGACGACGGGCTG 714

QY 141 IleProLeuArgValThrGluGluAlaGlyIleValAspPheThrSerValValSerGluIle 160
 DB 715 ATCCCCCTTAAGAGTCAACGAGAGGCTGAGAGGACTTACCTCTGTGTCTTCAGAAATC 774
 QY 161 MetMetTyrIleLeuLeuValPheLeuThrIleTyrPheLeuIleGluMetIleTyrCys 180
 DB 775 ATGATGTACATCTCTTGTGTCTTCTCTCACTGTGGCTCTCATGAGATGATACATTGC 834
 QY 181 -TyrArgIleValSerIleValGluGluAlaGlnGluAlaValAsnValSerAspTyrLeu 200
 DB 835 CTACAGAAACGGTATCAGACAGCAACGAGCCCAACGAAACCGAGATGCTACTTGG 894
 QY 200 1a1leProSerGluLeuValGluAsnSerAlaValPro 212
 DB 895 CGATTTCATTGAGAAACAGGAAATTTCTCGTACT 932
 RESULT 13
 AAS86764
 ID AAS86764 standard; cDNA; 1195 BP.
 XX
 AC AAS86764;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE DNA encoding novel human diagnostic protein #22568.
 XX
 KW Human: chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 PD 11-OCT-2001.
 XX
 XX 30-MAR-2001; 2001WO-US008631.
 XX
 PF 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 PI
 DR WPI; 2001-639362/73.
 DR P-PSDB; ABG22577.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 1; SEQ ID NO 22568; 103bp; English.
 XX
 XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probe, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS864197-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in

CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pat_sequences

XX Sequence 1195 BP; 253 A; 356 C; 381 G; 205 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 7,64e-101 Length: 1195
Score: 927.50 Matches: 184
Percent Similarity: 94.42% Conservative: 2
Best Local Similarity: 93.40% Mismatches: 9
Query Match: 82.52% Indels: 2
DB: 5 Gaps: 1

US-09-977-579-2 (1-215) x AAS86764 (1-1195)

```

QY 1 MetProAlaPheAsnArgLeuPheProLeuAlaSerLeuValLeuIleTyrTrpValSer 20
Db 463 ATGCGTGCCTTCATATGATTGTTTCCCTGGCTTCTCGGCTTATCTAGTCACTGAGTCACT 522
QY 21 ValCysPheProValCysValGluValProSerGluThrGluAla--ValGlnGlyAs 39
Db 523 GTCGCTTCCTCGTGGGAGAAAGTCCCTTAGAAAAAGCGGGCGCGGACCGGCTAA 582
QY 39 nProMetLeuAsnArgCysIleSerCysMetLeuArgGluGluValGluAlaThrThrVa 59
Db 583 CCCATGAAGCTGGCGTGCATCTCTGCATGAAGAGAGAGAGGCGGACACCGCT 642
QY 59 lValGluTrpPheTyrArgProGluGlyGlyLysAspPheLeuIleTyrGluTyrArgAs 79
Db 643 GGTGGAATGGTTCTTACAGAGCCGAGGCGGTAAGATTCTTATTATCGAGTATCGGAA 702
QY 79 nGlyHisGlnGluValGluSerProPheGlnGlyArgLeuGlnTrpAsnGlySerLysAs 99
Db 703 TGGCCACAGAGAGGTGAGAGCCCTTTCAGGGCGCTGAGTGAATGACAGCAAGCA 762
QY 99 pLeuGluAspValSerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuTyrTrh 119
Db 763 CCTGACGAGAGCTGTCATCACTGCTCAAGCTCACTGTAACGACTGGCTCTTAC 822
QY 119 rCysAsnValSerArgGluPheGluPheGluAlaHisArgProPheValLysThrThrAr 139
Db 823 CTGCAATGTGTCCGGAGGTTTGAAGTTCAGCGCATGGCCCTTTGTGAAGAGAGCGG 882
QY 139 gLeuIleProLeuArgValThrGluGluAlaGlyGluAspPheThrSerValValSerG 159
Db 883 GCTGATCCCTTAAGAGTCAACGAGAGGCTGAGAGACTTCACTCTGTGCTCAGA 942
QY 159 uIleMetMetTyrIleLeuLeuValPheLeuThrLeuTyrLeuLeuIleGluMetIleT 179
Db 943 AATCATGATGATTCCTTCTGTGTTCTTCTCACTTGTGGCTGCTCATCGAGATATTA 1002
QY 179 rCysTyrArgLysValSerLysAlaGluGluAlaAlaGlnLysAsnAla 195
Db 1003 TTGCTACGACAGGCTCTCAAAAGCCGAAGAGGACCCCAAAACGCG 1051

```

RESULT 14
ADSI0151
ID ADSI0151 standard; DNA; 953 BP.

AC ADI0151;
XX 16-DEC-2004 (first entry)
XX Human therapeutic DNA - SEQ ID 388.
XX antiinflammatory; neuroprotective; antianaemic; cytostatic; vulnerary;
XX inflammatory; haematopoiesis; immunity; neurodegenerative; stem cell;
XX aplastic anaemia; cancer; wound healing; gene therapy; ds; gene.
OS Homo sapiens.
XX WO2004080148-A2.
XX

PD 23-SEP-2004.
XX
XX 30-SEP-2003; 2003WO-US030720.
XX
XX 02-OCT-2002; 2002US-0416186P.
XX

PA (NUVE-) NUVELO INC.

XX Tang YT, Asundi V, Ren F, Zhang J, Wehrman T, Wang Z, Ma Y,
PI Wang D, Chen R, Zhao QA, Wang J, Ghosh M, Xue AJ, Weng G, Zhou P;
XX WPI: 2004-668857/65.
DR P-PSDB; ADSI0835.

PT New polynucleotide, useful in preparing a composition for diagnosing or
PT treating inflammatory, neurodegenerative or stem cell disorders, e.g.,
PT aplastic anemia or cancer for promoting wound healing.

PS Claim 1; SEQ ID NO 388; 718bp; English.

XX The invention relates to a novel isolated polynucleotide and the encoded
CC polypeptide. The molecules of the invention demonstrate antiinflammatory,
CC neuroprotective, antianaemic, cytostatic and vulnerary activities and may
CC be useful in preparing a composition for diagnosing or treating
CC inflammatory, haematopoietic, immune, neurodegenerative or stem cell
CC disorders, such as aplastic anaemia or cancer, as well as for promoting
CC wound healing. The molecules may also be utilised during gene therapy
CC procedures. The current sequence is that of a human therapeutic DNA of
CC the invention. The current sequence is not shown explicitly within the
XX specification but can be accessed from the WIPO web-site.

XX Sequence 953 BP; 180 A; 279 C; 290 G; 204 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4.53e-96 Length: 953
Score: 886.50 Matches: 174
Percent Similarity: 84.91% Conservative: 6
Best Local Similarity: 82.08% Mismatches: 19
Query Match: 78.87% Indels: 13
DB: 13 Gaps: 1

US-09-977-579-2 (1-215) x ADSI0151 (1-953)

```

QY 1 MetProAlaPheAsnArgLeuPheProLeuAlaSerLeuValLeuIleTyrTrpValSer 20
Db 327 ATGCGTGCCTTCATATGATTGTTTCCCTGGCTTCTCGGCTTATCTAGTCACTGAGTCACT 386
QY 21 ValCysPheProValCysValGluValProSerGluThrGluAlaGlnGlyAsnPro 40
Db 387 GTCGCTTCCTCGTGGTGTGAGAGTCCCTCGAGAGCGAGGCGGTCAGAGGCAACCCC 446
QY 41 MetLysLeuArgCysIleSerCysMetLysArgGluGluValGluAlaThrThrValVal 60
Db 447 ATGAGCTGGCGTGCATCTCTGCATGAAGAGAGAGAGTGAAGGCCACCAACGCTGTG 506
QY 61 GluTrpPheTyrArgProGluGlyGlyLysAspPheLeuIleTyrGluTyrArgAsnGly 80
Db 507 GAATGCTTCTTACAGGCCCGAGGCGGTAAAGATTCTTATTACGAGTATCGGAATGCG 566
QY 81 HisGlnGluValGluSerProPheGlnGlyArgLeuGlnTrpAsnGlySerLysAspLeu 100
Db 567 CACCGAGAGGTGAGAGCCCTTTCAGGGCGCTGCAAGAGAGAGAGAGAGAGACCTG 626
QY 101 GlnAspValSerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuTyrTrhCys 120
Db 627 CAGAGCTGTGCATCACTGCTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 686
QY 121 AsnValSerArgGluPheGluPheGluAlaHisArgProPheValLysThrThrArgLeu 140
Db 687 AATGTGTCCCGGAGTTTGAAGTTCAGGCGCATCGGCCCTTGTGAGAGAGAGAGCGGCTG 746
QY 141 lLeuProLeuArgValThrGluGluAlaGlyGluAspPheThrSerValValSerGluIle 160

```

Db 747 ATCCCTTAAGTACACCGAGAGCTGAGAGACTTCACTCTGTGTCTAACACTTC 806
 QY 161 MetMetTyrTlleuLeuValPheLeuThrLeuTyrLeuLeuIleGluMetIleTyrCys 180
 Db 807 ATGATGACATCCTTCTGTGTCTTCCACCTT----- 839
 QY 181 TyrArgValSerLysAlaGluGluAlaGlnGluAsnAlaSerAspTyrLeuAla 200
 Db 840 -----GTGAACCAACGCCACCGCTCCAGAGCAGCATCTACTGCTCCACCACT 887
 QY 201 IleProSerGluAsnLysGluAsnSerAlaValPro 212
 Db 888 GCTCTGATGATCATCTGGGCTTCCAGCATCCCT 923
 RESULT 15
 AAV86895
 ID AAV86895 standard; cDNA; 471 BP.
 XX
 AC AAV86895;
 XX
 DT 27-APR-1999 (first entry)
 XX
 DE EST clone BM4.
 XX
 KM Expressed sequence tag; secreted protein; haematopoiesis regulator;
 KM tissue growth; activating; inhibin; tumour invasion suppressor; EST; human;
 KM chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;
 KM receptor; ligand; anti-inflammatory; tumour inhibitor; ds.
 XX
 OS Homo sapiens.
 XX
 PN MO845435-A2.
 XX
 PD 15-OCT-1998.
 XX
 PF 10-APR-1998; 98MO-US006954.
 XX
 PR 10-APR-1997; 97US-00835913.
 XX
 PA (GEMV) GENETICS INST INC.
 XX
 PI Jacobs K, McCoy JM, Lavallie ER, Racie LA, Werberg D, Treacy M;
 PI Spaulding V, Agostino MJ;
 XX
 PI human blood, kidney, foetal lung, placenta, testes, brain, ovary,
 PT pituitary, retina and colon cDNA libraries.
 PT
 XX
 DR WPI; 1999-070076/06.
 XX
 PS Claim 1; Page 383-384; 633pp; English.
 XX
 CC This sequence represents an expressed sequence tag (EST), and is a
 CC polynucleotide of the invention. The polynucleotides of the invention are
 CC all secreted EST sequences isolated from a variety of human tissue
 CC sources. The EST sequences and proteins encoded by them are predicted to
 CC have useful biological activities which would make them suitable for
 CC treating, preventing or ameliorating medical conditions in humans and
 CC animals, although no supporting data is given. Suggested activities
 CC include nutritional activity, immune stimulating or suppressing activity,
 CC haematopoiesis regulating activity, tissue growth activity,
 CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
 CC activity, cadherin/tumour invasion suppressor activity, tumour inhibition
 CC activity. The EST sequences are also stated to be useful for gene therapy
 XX
 SQ Sequence 471 BP; 93 A; 126 C; 143 G; 108 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.: 3,44e-79 Length: 471
 Score: 742.00 Matches: 143
 Percent Similarity: 98.62% Conservative: 0
 Best Local Similarity: 98.62% Mismatches: 2

Query Match: 66.01% Indels: 1
 DB: 2 Gaps: 0
 US-09-977-579-2 (1-215) x AAV86895 (1-471)
 QY 1 MetProAlaPheAsnArgLeuPheProLeuAlaSerLeuValLeuIleTyrTrpValSer 20
 Db 29 ATGCGTCCCTTAATGATGATTGTTCCCTGGGTTCTCTGTGCTTATCTACGGGTCACT 88
 QY 21 ValCysPheProValCysValGluValProSerGluThrGluAlaValGlnGlyAsnPro 40
 Db 89 GTCTGCTTCCCTGTGTGTGTGGAAAGCCCTCGAGACGAGCC-GTGCANGGCAACCCC 147
 QY 41 MetLeuLeuArgCysIleSerCysMetLysArgGluGluValGluAlaThrThrValVal 60
 Db 148 ATGAACTCGCGCTGCATCTCTGCAATGAAGAGAGAGAGAGGACCAACAGTGTG 207
 QY 61 GluTrpPheTyrArgProGluGlyGlyLysAspPheLeuIleTyrGluTyrArgAsnGly 80
 Db 208 GAATGTTCTACAGGCCCGGCGGCGGTAAAGATTCTTATTACGAGTATCGAATGCG 267
 QY 81 HisGlnGluValGluSerProPheGlnGlyArgLeuGlnTrpAsnGlySerLysAspLeu 100
 Db 268 CACCAAGAGGTGAGAGCCCTTTCAGGGCCGCTGCAAGTGAATGGCAGCAAGACCTG 327
 QY 101 GlnAspValSerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuTyrThrCys 120
 Db 328 CAGAGCTGTCCATCTACTGTCTCAACGTCATCTGAAAGACTGTGCTCTACACTGCG 387
 QY 121 AsnValSerArgGluPheGluPheGluAlaHisArgProPheValLysThrThrArgLeu 140
 Db 388 AATGTGTCCCGGAGATTGAGTTGAGCGGCATCGCCCTTGTGAAAGACGACGCGCTG 447
 QY 141 IleProLeuArgVal 145
 Db 448 ATCCCTTAAGAGTC 462

Search completed: April 7, 2005, 15:26:59
 Job time : 437 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 7, 2005, 15:00:45 ; Search time 162 Seconds
(without alignments)

2171.602 Million cell updates/sec

Title: US-09-977-579-2

Perfect score: 1124

Sequence: 1 MPAPNRLFLPLASLVLIYVWS.....SDYLAIPSENKENSAPVVEE 215

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09977579/runat_06042005_163457_23620/app_query.fasta_1.391
-DB=Issued Patents NA -Qfmt=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USR=US09977579@cgn_1.1.105@runat_06042005_163457_23620 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 470 | 41.8 | 617 | 4 | US-09-949-016-2369 |
| 2 | 165.5 | 14.7 | 11174 | 4 | Sequence 2369, Ap |
| 3 | 165.5 | 14.7 | 30337 | 4 | Sequence 14111, A |
| 4 | 147 | 13.1 | 1342 | 4 | Sequence 12053, A |
| 5 | 147 | 13.1 | 1708 | 4 | Sequence 787, App |
| 6 | 131 | 11.7 | 1121 | 4 | Sequence 4948, Ap |
| 7 | 131 | 11.7 | 1151 | 3 | Sequence 5103, Ap |
| 8 | 131 | 11.7 | 1151 | 3 | Sequence 1, Appli |
| 9 | 131 | 11.7 | 1151 | 3 | Sequence 3, Appli |
| 10 | 131 | 11.7 | 1151 | 3 | Sequence 5, Appli |
| 11 | 129 | 11.5 | 600 | 3 | Sequence 7, Appli |
| 12 | 129 | 11.5 | 600 | 3 | Sequence 43, Appli |
| | | | | | Sequence 47, Appli |

| | | | | | | |
|----|-------|------|------|---|--------------------|-------------------|
| 13 | 127 | 11.3 | 600 | 3 | US-09-430-503-41 | Sequence 41, Appl |
| 14 | 127 | 11.3 | 600 | 3 | US-09-430-503-45 | Sequence 45, Appl |
| 15 | 127 | 11.3 | 752 | 3 | US-09-430-503-17 | Sequence 17, Appl |
| 16 | 127 | 11.3 | 752 | 3 | US-09-430-503-19 | Sequence 19, Appl |
| 17 | 127 | 11.3 | 752 | 3 | US-09-430-503-21 | Sequence 21, Appl |
| 18 | 127 | 11.3 | 752 | 3 | US-09-430-503-23 | Sequence 23, Appl |
| 19 | 125 | 11.1 | 1073 | 3 | US-09-430-503-25 | Sequence 25, Appl |
| 20 | 125 | 11.1 | 1073 | 3 | US-09-430-503-27 | Sequence 27, Appl |
| 21 | 125 | 11.1 | 1073 | 3 | US-09-430-503-29 | Sequence 29, Appl |
| 22 | 125 | 11.1 | 1073 | 3 | US-09-430-503-31 | Sequence 31, Appl |
| 23 | 124.5 | 11.1 | 1060 | 4 | US-09-799-451-222 | Sequence 222, App |
| 24 | 118 | 10.5 | 480 | 3 | US-09-430-503-33 | Sequence 33, Appl |
| 25 | 118 | 10.5 | 480 | 3 | US-09-430-503-35 | Sequence 35, Appl |
| 26 | 118 | 10.5 | 480 | 3 | US-09-430-503-37 | Sequence 37, Appl |
| 27 | 118 | 10.5 | 480 | 3 | US-09-430-503-39 | Sequence 39, Appl |
| 28 | 112.5 | 10.0 | 1592 | 3 | US-09-656-952-1 | Sequence 1, Appli |
| 29 | 110.5 | 9.8 | 1111 | 4 | US-09-949-016-4949 | Sequence 4949, Ap |
| 30 | 110.5 | 9.8 | 1113 | 4 | US-09-949-016-652 | Sequence 652, App |
| 31 | 109 | 9.7 | 1207 | 4 | US-09-023-655-131 | Sequence 131, App |
| 32 | 109 | 9.7 | 1561 | 3 | US-09-656-952-18 | Sequence 18, Appl |
| 33 | 109 | 9.7 | 1591 | 3 | US-09-656-952-17 | Sequence 17, Appl |
| 34 | 109 | 9.7 | 1751 | 4 | US-09-855-323-11 | Sequence 11, Appl |
| 35 | 108.5 | 9.7 | 2184 | 3 | US-09-484-970B-161 | Sequence 161, App |
| 36 | 108 | 9.6 | 8083 | 3 | US-09-383-630-4 | Sequence 4, Appli |
| 37 | 108 | 9.6 | 8083 | 3 | US-09-383-630-5 | Sequence 5, Appli |
| 38 | 106 | 9.4 | 1557 | 3 | US-08-996-338-2 | Sequence 2, Appli |
| 39 | 106 | 9.4 | 1557 | 4 | US-09-556-972-2 | Sequence 2, Appli |
| 40 | 106 | 9.4 | 2830 | 1 | US-08-604-333-3 | Sequence 3, Appli |
| 41 | 106 | 9.4 | 2830 | 3 | US-09-110-618-3 | Sequence 3, Appli |
| 42 | 106 | 9.4 | 2830 | 3 | US-09-578-178-3 | Sequence 3, Appli |
| 43 | 106 | 9.4 | 2830 | 4 | US-09-577-806-3 | Sequence 3, Appli |
| 44 | 106 | 9.4 | 2830 | 4 | US-09-621-502-7 | Sequence 7, Appli |
| 45 | 99.5 | 8.9 | 2079 | 4 | US-09-949-016-3171 | Sequence 3171, Ap |

ALIGNMENTS

RESULT 1

US-09-949-016-2369
; Sequence 2369, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2369
; LENGTH: 617
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-2369

Alignment Scores:
Pred. No.: 1.36e-53 Length: 617
Score: 470.00 Matches: 101
Percent Similarity: 66.16% Conservative: 30
Best Local Similarity: 51.01% Mismatches: 59
Query Match: 41.81% Indels: 8
DB: 4 Gaps: 5

US-09-977-579-2 (1-215) x US-09-949-016-2369 (1-617)

Qy 26 CyaValgluValProSerGluThrGluAlaValGlnGlyAsnProMetLysLeuArgCys 45

```

Db      21  TGGTGGAGGTGACTCGGAGCCGCGGTGATGGATGACCTTCAAAATCTTTCG 80
QY      46  ILeSerCysMetLysArgGluGluValGluAlaThrThrValValGluTrpPheTyrArg 65
Db      81  ATCTCTGCAAGCGCGGAGGAGCAACAGCTGAGACCTTACCGAGTGGACCTTCGCG 140
QY      66  ProGluGlyGlyLysAspPheLeu---IleTyrGluTyrArgAsnGlyHisGlnGluVal 84
Db      141  CAGAAGGCACGTGAGGAGTTGTCAAGATCTCGCGCTATGAGAATGAGGTGTGACAGTG 200
QY      85  GluSerPro-----PheGlnGlyArgLeuGlnTrpAsnGlySer-----LysAsp 99
Db      201  GAGGAGGATGAGCGCTTCGAGGCGCGGTGGTGGAAATGGCAGCGCGGCCACCAAGAC 260
QY      100  LeuGlnAspValSerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuTyrThr 119
Db      261  CTGCAGGATCTGTATCTTATCATCCATGTCACCTACACCACTCGGGCGACTACGAG 320
QY      120  CysAsnValSerArgGluPheGluPheGluAlaHisArgProPheValLysThrThrArg 139
Db      321  TGCACGCTACCGCTGCTCTTCTTCGAAAACTACGAGCACAACACCAAGCGTCTGTCAG 380
QY      140  LeuIleProLeuArgValThrGluGluAlaGlyGluAspPheThrSerValValSerGlu 159
Db      381  AAGATCCACATTGAGTGTGCAAGAACCAAGACAGACATGGCATCCATCGTGTCTGAG 440
QY      160  IleMetMetTyrIleLeuLeuValPheLeuThrLeuTrpLeuLeuGluMetIleTyr 179
Db      441  ATCATGATGATGTGCTCATTTGTTGGTGGTACCATATGCTGTCGCGCAGATGATTAC 500
QY      180  CysTyrArgLysValSerLysAla---GluGluAlaAlaGlnGluAsnAlaSerAspTyr 198
Db      501  TGCTACAGAAGATCGCTCGCGCAGGAGACTGCTGCACAGAGAAATCGCTCGGAATAC 560
QY      199  LeuAlaIleProSerGluAsnLysGluAsn---SerAlaValProValGluGlu 215
Db      561  CTGGCCATCACCTCTGAAAGAAAGAAAGAACTGCACGCGGCTCAGGTGGCCGAA 614

```

RESULT 2

```

US-09-949-016-14111
; Sequence 14111, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14111
; LENGTH: 11174
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(11174)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14111

```

Alignment Scores:
 Pred. No.: 1.5e-10 Length: 11174
 Score: 165.50 Matches: 56
 Percent Similarity: 47.13% Conservative: 26
 Best Local Similarity: 32.18% Mismatches: 45

```

Query Match: 14.72% Indels: 48
DB: 4 Gaps: 7
US-09-977-579-2 (1-215) x US-09-949-016-14111 (1-11174)
QY      67  GluGlyGlyLysAspPheLeuIleTyrGlu-TyrArgAsnGlyHisGlnGluValGluSe 86
Db      8220  GAAGAGGTATTTCATTTTACAAATTGAAGAAACAGAGG-----CCTAGAGACATTGAGTC 8273
QY      86  rPropheGlnGlyArgLeu-----GlnTrp-----AsnGlySerLysAs 99
Db      8274  ACTTCGCAAGGTACACAGCTGGCGAGTGGCAGAGCCAGCTTTGAACCTGAGCCTTCT 8333
QY      99  pleuGlnAspValSerIle-----ThrValLeuAsnValThrLeuAsnAspSerGlyLe 117
Db      8334  GGTTCAGAGCTACCGCTTTTAGGCACCGTGTGAGGGCGCTCCAG-AATGACACAGATGT 8392
QY      117  uTyrThrCysAsnValSerArgGluPheGluPheGluAlaHisArg----- 132
Db      8393  GGCCTCGAGTTACACAGGAGCGAGTTGAGGGTGACCCCGCCAGGCGCTACCAAGGCT 8452
QY      133  -----ProPheValLysThrThrArgLeuIleProLeu----- 143
Db      8453  GGGTATTAAATACACAGTGCATACACAGCGCCAGGAGGTTGAGCCACTCATCCAAGC 8512
QY      143  ----- 143
Db      8513  TCACACAGCAAGCTCACACACACTCAGGCTGTCTATGAGCGCTGGGCTACCCCTTAACC 8572
QY      144  -ArgValThrGluGluAlaGlyGluAspPheThrSerValValSerGluIleMetMetTy 163
Db      8573  CTGCTGCGCCCTGCGCCACACAGACATGGCATCCATCGTGTCTGAGATCATGATGA 8632
QY      163  rIleLeuLeuValPheLeuThrLeuTrpLeuLeuGluMetIleTyrCysTyrArgLy 183
Db      8633  TGTGCTCATTTGTGGTGTGACCATATGCTGTCGCGCAGAGATGATTTACTGCTACAGAA 8692
QY      183  sValSerLysAla---GluGluAlaAlaGlnGluAsnAla 195
Db      8693  GATCGCTCGCGCCAGGAGACTGCTGCACAGAGAATCGG 8732

```

RESULT 3

```

US-09-949-016-13053
; Sequence 13053, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13053
; LENGTH: 30337
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(30337)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13053

```

Alignment Scores:
 Pred. No.: 7.24e-10 Length: 30337
 Score: 165.50 Matches: 56


```
; LENGTH: 1708
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-4948

Alignment Scores:
Pred. No.:          Length:          1708
Score:             Matches:          57
Percent Similarity: 41.44%    Conservative: 35
Best Local Similarity: 25.68%    Mismatches: 80
Query Match:       13.08%    Indels: 50
DB:                4        Gaps: 9

US-09-977-579-2 (1-215) x US-09-949-016-4948 (1-1708)

QY 7 LeuPheProLeuAlaSerLeuValLeuIleTyrTrpValSerValCysPheProValCys 26
Db 225 CTTTGGCCTATAGCAGCT---GTGGAAATTTATACCTCCCGGGTG-----266
QY 27 ValGluValProSerGluThrGluAlaValGlnGlyAsnProMetLysLeuArgCysIle 46
Db 267 -----CTGGAGGCTGTAAATGGGACAGATGCTCGGTAAATGCACT 308
QY 47 SerCysMetLysArgGluGluValGluAlaThrThrValValGluTrpPheTyrArgPro 66
Db 309 TTCTCAGCTTTGCCCTGTGGGTGATGCTCTAACAGTGACC---TGGAAATTTTCGTCT 365
QY 67 ---GluGlyGly---LysAspPheLeuIleTyrGluTyrArgAsnGlyHisGlnGluVal 84
Db 366 CTAGACGGGGGACCTGAGCAGTTGTATTCTACTACACATAGATCCCTTCCACCCATG 425
QY 85 GluSerProPheGlnGlyArgLeuGlnTrpAsnGlySerLysAspLeuGlnAspValSer 104
Db 426 AGTGGCGGTTTAAAGCACCGGTTCTTGGGATGGGAATCCTGAGCGGTACGATGCTCCT 485
QY 105 IleThrValLeuAsnValThrLeuAsnAspSerGlyLeuTyrThrCysAsnValSerArg 124
Db 486 ATCCTTCTCTGGAAACTCGAGTCGACGACAAATGGGACATACACCTGCCAGGTGAAGAC 545
QY 125 GluPheGluPheGluAlaHisArgProPheValLysThrThrArgLeuIleProLeuArg 144
Db 546 -----CCACCTGAT 554
QY 145 ValThrGluGluAlaGlyGluAspPheThrSerValVal-----SerGlu 159
Db 555 GTTGATGGGGTGATAGGGGAGATCCGGCTCGCTGCACACTGTACGCTTCTCTGAG 614
QY 160 IleMetMetTyrIleLeuLeuVal-----PheLeuThrLeuTrpLeu 173
Db 615 ATCCACTTCTGGCTCTGCCCATTTGGCTCTGCTGTGCACTGATGATCATATAGTAAT 674
QY 174 LeuIleGluMetIleTyrCysTyrArgLysValSerLysAlaGluAlaGlnGlu 193
Db 675 GTAGTGGTCTCTTCCAGCATTACCGGAAAAAGCGATGGCGCGAAAGACT-----725
QY 194 AsnAlaSerAspTyrLeuAlaIleProSerGluAsnLysGluAsnSerAlaValProVal 213
Db 726 -----CATAAAGTGGTGGGATATAAATCAAAAGAGAGAAAGGCTCAACCAA 773
QY 214 GluGlu 215
Db 774 GAGAAA 779

RESULT 6
US-09-949-016-5103
; Sequence 5103, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
```

```
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5103
; LENGTH: 1121
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-5103

Alignment Scores:
Pred. No.:          Length:          1121
Score:             Matches:          51
Percent Similarity: 37.39%    Conservative: 32
Best Local Similarity: 22.97%    Mismatches: 81
Query Match:       11.65%    Indels: 58
DB:                4        Gaps: 9

US-09-977-579-2 (1-215) x US-09-949-016-5103 (1-1121)

QY 27 ValGluValProSerGluThrGluAlaValGlnGlyAsnProMetLysLeuArgCysIle 46
Db 250 GTATATACGCCAAAAGAAATCTTCGTGCAAAATGGTACACAAAGGGAAGCTGACCTGC---306
QY 47 SerCysMetLysArgGluGluValGluAlaThr-----ThrValValGluTrpPhe 63
Db 307 -----AGTTCAAGTCTACTAGTAGCAGCTGGCGGGTTCACCTCAGTCTCTGGAGC 357
QY 64 TyrArgProGluGlyGlyLysAspPheLeu---IleTyrGluTyrArgAsnGlyHisGln 82
Db 358 TTCACGACGAGGGGGCGGACACTACTGTGTCGTTTTTCCACTACTCCCAAGGGCAAGTG 417
QY 83 GluValGluSer-----ProPheGlnGlyArgLeuGlnTrpAsnGlySerLysAspLeu 100
Db 418 TACCTTGGGAATTAFCACCAATTTAAAGACAGAAATCAGCTGGGCTGGAGACCTTACACAG 477
QY 101 GlnAspValSerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuTyrThrCys 120
Db 478 AAAGATGATCATCATCAACATAGAAATATGAGTTTATACACATGGCACCTATATCTGT 537
QY 121 AsnValSerArgGluPheGluPheGluAlaHisArgProPheValLysThrThrArgLeu 140
Db 538 GATGTCAAAAAC-----549
QY 141 IleProLeuArgValThrGluGluAlaGlyGluAspPheThrSerValValSerGlu---159
Db 550 ---CCTCTGACATCGTTGTCGAGCTCGACACATAGGCTCTATGTCTGTAAGAAAGAG 606
QY 160 -----IleMetMetTyrIleLeuLeuValPheLeuThr 170
Db 607 AATTTGCTGTGTTTCCAGTTTGGTAGTGGTGGGACATAGTTACTGCTGTGCTCTAGCT 666
QY 171 LeuTrpLeuLeuIleGluMetIleTyrCys-----TyrArgLys-----183
Db 667 CTCACCTGCTCATCAGCATGATTCTGCTGCTCTATAGAAGGAAAAAAGCTCTAAACGG 726
QY 184 -----ValSerLysAlaGluGluAlaGlnGlu 193
Db 727 GATTACACTGGCTGCAGTACATCAGAGAGTTTGTCCACCATTTAGCAGCTCTCTCGAAG 786
QY 194 AsnAlaSerAsp-----TyrLeuAlaIleProSerGluAsnLysGluAsnSer 209
Db 787 TCCCCCTCCGACACTGAGGGTCTTTGAAAGAGTCTGCTCTCTCTGATCTCTCACCAGGGCCCA 846
QY 210 AlaVal 211
Db 847 GTCAT 852
```

```
RESULT 7
US-09-430-503-1
; Sequence 1, Application US/09430503
; Patent No. 6355786
; GENERAL INFORMATION:
; APPLICANT: Zhao, Zhizhuang
; TITLE OF INVENTION: PURIFIED AND ISOLATED PROTEIN ZERO RELATED (PZR) AND
; TITLE OF INVENTION: THERAPEUTIC AND SCREENING METHODS USING SAME
; FILE REFERENCE: Attorney Docket No. 6355786 1242-11/2
; CURRENT APPLICATION NUMBER: US/09/430,503
; CURRENT FILING DATE: 1999-10-29
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1151
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: gene
; LOCATION: (1)..(1151)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (133)..(939)
US-09-430-503-1
Alignment Scores:
Pred. No.: 1.85e-07 Length: 1151
Score: 131.00 Matches: 51
Percent Similarity: 37.39% Conservative: 32
Best Local Similarity: 22.97% Mismatches: 81
Query Match: 11.65% Indels: 58
DB: Gaps: 9

US-09-977-579-2 (1-215) x US-09-430-503-1 (1-1151)
QY 27 ValGluValProSerGluThrGluAlaValGlnGlyAsnProMetLysLeuArgCysIle 46
DB 250 GTATATACGCCAAAGAAATCTTCGTGGCAAAATGGTACACAGGGAAGCTGACCTGC--- 306
QY 47 SerCysMetLysArgGluGluValGluAlaThr-----ThrValValGluTrpPhe 63
DB 307 -----AAGTTCAGTCTACTAGTACGACTGGCGGGTTGACCTCCTCTCGAGC 357
QY 64 TyrArgProGluGlyLysAspPheLeu---IleTyrGluTyrArgAsnGlyHisGln 82
DB 358 TTCAGCCAGAGGGCGGCACACTACTGTCTGCTTTTCCACTACTCCCAAGGCAAGTG 417
QY 83 GluValGluSer-----ProPheGlnGlyArgLeuGlnTrpAsnGlySerLysAspLeu 100
DB 418 TACCTTGGGAATATCCACCATTTAAAGACAGAATCAGCTGGCTGGAGACCTTGACAAG 477
QY 101 GluAspValSerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuTyrThrCys 120
DB 478 AAGATGTCATCAATCAACATAGAAAATATGACGTTTATACACAAATGGCACCCTATATCTGT 537
QY 121 AsnValSerArgGluPheGluAlaHisArgProPheValLysThrArgLeu 140
DB 538 GATGTCAAAAAC-----ValSerLysAlaGluGluAlaGlnGlu 193
QY 141 IleProLeuArgValThrGluGluAlaGlyGluAspPheThrSerValValSerGlu--- 159
DB 550 ---CCTCCTGACATCGTTGTCAGCTGGTGGGATAGTTACTGCTGTGGTCTTAGGT 666
QY 160 -----IleMetMetTyrIleLeuLeuValPheLeuThr 170
DB 607 AATTGCTGTGTTTCAGTTTGGTAGTGGTGGGATAGTTACTGCTGTGGTCTTAGGT 666
QY 171 LeuTrpLeuLeuIleGluMetIleTyrCys-----TyrArgLys 183
DB 667 CTACTCTGCTCATCAGCATGATTCGTGCTGCTCTATAGAGGAAAAAACTCTAAACGG 726
QY 184 -----ValSerLysAlaGluGluAlaGlnGlu 193
DB 184 -----ValSerLysAlaGluGluAlaGlnGlu 193

RESULT 8
US-09-430-503-3
; Sequence 3, Application US/09430503
; Patent No. 6355786
; GENERAL INFORMATION:
; APPLICANT: Zhao, Zhizhuang
; TITLE OF INVENTION: PURIFIED AND ISOLATED PROTEIN ZERO RELATED (PZR) AND
; TITLE OF INVENTION: THERAPEUTIC AND SCREENING METHODS USING SAME
; FILE REFERENCE: Attorney Docket No. 6355786 1242-11/2
; CURRENT APPLICATION NUMBER: US/09/430,503
; CURRENT FILING DATE: 1999-10-29
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1151
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: gene
; LOCATION: (1)..(1151)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (133)..(939)
US-09-430-503-3
Alignment Scores:
Pred. No.: 1.85e-07 Length: 1151
Score: 131.00 Matches: 51
Percent Similarity: 37.39% Conservative: 32
Best Local Similarity: 22.97% Mismatches: 81
Query Match: 11.65% Indels: 58
DB: Gaps: 9

US-09-977-579-2 (1-215) x US-09-430-503-3 (1-1151)
QY 27 ValGluValProSerGluThrGluAlaValGlnGlyAsnProMetLysLeuArgCysIle 46
DB 250 GTATATACGCCAAAGAAATCTTCGTGGCAAAATGGTACACAGGGAAGCTGACCTGC--- 306
QY 47 SerCysMetLysArgGluGluValGluAlaThr-----ThrValValGluTrpPhe 63
DB 307 -----AAGTTCAGTCTACTAGTACGACTGGCGGGTTGACCTCCTCTCGAGC 357
QY 64 TyrArgProGluGlyLysAspPheLeu---IleTyrGluTyrArgAsnGlyHisGln 82
DB 358 TTCAGCCAGAGGGCGGCACACTACTGTCTGCTTTTCCACTACTCCCAAGGCAAGTG 417
QY 83 GluValGluSer-----ProPheGlnGlyArgLeuGlnTrpAsnGlySerLysAspLeu 100
DB 418 TACCTTGGGAATATCCACCATTTAAAGACAGAATCAGCTGGCTGGAGACCTTGACAAG 477
QY 101 GluAspValSerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuTyrThrCys 120
DB 478 AAGATGTCATCAATCAACATAGAAAATATGACGTTTATACACAAATGGCACCCTATATCTGT 537
QY 121 AsnValSerArgGluPheGluAlaHisArgProPheValLysThrArgLeu 140
DB 538 GATGTCAAAAAC-----ValSerLysAlaGluGluAlaGlnGlu 193
QY 141 IleProLeuArgValThrGluGluAlaGlyGluAspPheThrSerValValSerGlu--- 159
DB 550 ---CCTCCTGACATCGTTGTCAGCTGGTGGGATAGTTACTGCTGTGGTCTTAGGT 666
QY 160 -----IleMetMetTyrIleLeuLeuValPheLeuThr 170
DB 607 AATTGCTGTGTTTCAGTTTGGTAGTGGTGGGATAGTTACTGCTGTGGTCTTAGGT 666
QY 171 LeuTrpLeuLeuIleGluMetIleTyrCys-----TyrArgLys 183
DB 667 CTACTCTGCTCATCAGCATGATTCGTGCTGCTCTATAGAGGAAAAAACTCTAAACGG 726
QY 184 -----ValSerLysAlaGluGluAlaGlnGlu 193
DB 184 -----ValSerLysAlaGluGluAlaGlnGlu 193
```

QY 160 -----lleMetMetTyrIleLeuLeuValPheLeuThr 170
Db 607 AATTGCTGCTGTTCCAGTTGGGTAGTGGGCATAGTTACTGCTGCTGCTAGGT 666
QY 171 LeuTrpLeuLeuLeuMetIleTyrCys-----TyrArgLys----- 183
Db 667 CTCACCTGCTCATCAGCATGATTCTGGCTGCTCTATAGAGGAAACTCTTAACGG 726
QY 184 -----ValSerLysAlaGluGluAlaAlaGluGlu 193
Db 727 GATTACACTGGCTGCAGTACATCAGAGAGTTGTCCACAGTTAAGCAGGCTCTCGGAG 786
QY 194 AsnAlaSerAsp-----TyrLeuAlaIleProSerGluAsnLysGluAsnSer 209
Db 787 TCCCTCCGACACTGAGGCTCTTGTAAAGAGTCTGCTTCTGGATCTCACCAGGGCCCA 846
QY 210 AlaVal 211
Db 847 GTCATA 852
RESULT 9
US-09-430-503-5
; Sequence 5, Application US/09430503
; Patent No. 6355786
; GENERAL INFORMATION:
; APPLICANT: Zhao, Zhizhuang
; TITLE OF INVENTION: PURIFIED AND ISOLATED PROTEIN ZERO RELATED (PZR) AND
; FILE REFERENCE: ATTORNEY DOCKET NO. 6355786 1242-11/2
; CURRENT APPLICATION NUMBER: US/09/430,503
; CURRENT FILING DATE: 1999-10-29
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1151
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: gene
; LOCATION: (1)..(1151)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (133)..(939)
US-09-430-503-5
Alignment Scores:
Pred. No.: 1.85e-07 Length: 1151
Score: 131.00 Matches: 51
Percent Similarity: 37.39% Conservative: 32
Best Local Similarity: 22.97% Mismatches: 81
Query Match: 11.65% Indels: 58
DB: 3 Gaps: 9
US-09-977-579-2 (1-215) x US-09-430-503-5 (1-1151)
QY 27 ValGluValProSerGluThrGluAlaValGlnGlyAsnProMetLysLeuArgCysIle 46
Db 250 GTATATACGCCAAAGAAATCTTCGTGGCAAAATGGTACACAAAGGGAAGCTGACCTGC--- 306
QY 47 SerCysMetLysArgGluValGluAlaThr-----ThrValValGluTrpPhe 63
Db 307 -----AAGTTCAGTCTACTAGTACGACTGGGGGTTCACCTCAGTCTCTCTGGAGC 357
QY 64 TyrArgProGluGlyCysLysAspPheLeu-----lleTyrGluTyrArgAsnGlyHisGln 82
Db 358 TTCACGACGAGAGGGGGCCGACACTACTGTGCTGCTTTTCCACTACTCTCCCAAGGCAAGTG 417
QY 83 GluValGluSer-----ProPheGlnGlyArgLeuGlnTrpAsnGlySerLysAspLeu 100
Db 418 TACCTTGGGAATATCCACCATTTAAAGACAGAAATCAGCTGGGCTGGAGACCTTGACACAG 477
QY 101 GlnAspValSerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuTyrThrCys 120
Db -----AAGTTCAGTCTACTAGTACGACTGGGGGTTCACCTCAGTCTCTCTGGAGC 357

Db 478 AAGATGTCATCATCAATACATAGAAAATATGAGTTTATACAAATGGCACCTATATCTGT 537
QY 121 AsnValSerArgGluPheGluPheGluAlaHisArgProPheValLysThrArgLeu 140
Db 538 GATGTCAAAAAC----- 549
QY 141 IleProLeuArgValThrGluGluAlaGlyGluAspPheThrSerValValSerGlu--- 159
Db 550 ---CCTCTGCATCGTTGTCAGCTGCACACATAGGCTCTATGCTGTAAGAAAGAG 606
QY 160 -----lleMetMetTyrIleLeuLeuValPheLeuThr 170
Db 607 AATTGCTGCTGTTTCCAGTTGGGTAGTGGGCATAGTTACTGCTGCTGCTAGGT 666
QY 171 LeuTrpLeuLeuLeuMetIleTyrCys-----TyrArgLys----- 183
Db 667 CTCACCTGCTCATCAGCATGATTCTGGCTGCTCTATAGAGGAAACTCTTAACGG 726
QY 184 -----ValSerLysAlaGluGluAlaAlaGluGlu 193
Db 727 GATTACACTGGCTGCAGTACATCAGAGAGTTGTCCACAGTTAAGCAGGCTCTCGGAG 786
QY 194 AsnAlaSerAsp-----TyrLeuAlaIleProSerGluAsnLysGluAsnSer 209
Db 787 TCCCTCCGACACTGAGGCTCTTGTAAAGAGTCTGCTTCTGGATCTCACCAGGGCCCA 846
QY 210 AlaVal 211
Db 847 GTCATA 852
RESULT 10
US-09-430-503-7
; Sequence 7, Application US/09430503
; Patent No. 6355786
; GENERAL INFORMATION:
; APPLICANT: Zhao, Zhizhuang
; TITLE OF INVENTION: PURIFIED AND ISOLATED PROTEIN ZERO RELATED (PZR) AND
; FILE REFERENCE: ATTORNEY DOCKET NO. 6355786 1242-11/2
; CURRENT APPLICATION NUMBER: US/09/430,503
; CURRENT FILING DATE: 1999-10-29
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1151
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: gene
; LOCATION: (1)..(1151)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (133)..(939)
US-09-430-503-7
Alignment Scores:
Pred. No.: 1.85e-07 Length: 1151
Score: 131.00 Matches: 51
Percent Similarity: 37.39% Conservative: 32
Best Local Similarity: 22.97% Mismatches: 81
Query Match: 11.65% Indels: 58
DB: 3 Gaps: 9
US-09-977-579-2 (1-215) x US-09-430-503-7 (1-1151)
QY 27 ValGluValProSerGluThrGluAlaValGlnGlyAsnProMetLysLeuArgCysIle 46
Db 250 GTATATACGCCAAAGAAATCTTCGTGGCAAAATGGTACACAAAGGGAAGCTGACCTGC--- 306
QY 47 SerCysMetLysArgGluValGluAlaThr-----ThrValValGluTrpPhe 63
Db 307 -----AAGTTCAGTCTACTAGTACGACTGGGGGTTCACCTCAGTCTCTCTGGAGC 357

Db 118 GTATATACGCCAAAGAAATCTTCGTGGCAATGGTACACAAGGGAGCTGACCTGC--- 174
QY 47 SerCysMetLysArgGluGluValGluAlaThr-----ThrValValGluTrpPhe 63
Db 175 -----AAGTTCAAGTCTACTAGTACGACTGGGGGTTGACCTCAGCTCTCTCGGAGC 225
QY 64 TyrArgProGluGlyGlyLysAspPheLeu---IleTyrGluTyrArgAsnGlyHisGln 82
Db 226 TTCCAGCAGAGGGGGCCGACACTACTGTGTGCTTTTCCACTACTCCCAAGGGCAAGTG 285
QY 83 GluValGluSer-----ProPheGlnGlyArgLeuGlnTrpAsnGlySerLysAspLeu 100
Db 286 TACATTGGGAATATCCACCAATTTAAAGACAGAGATCAGCTGGCTGGAGACCTTGACAAG 345
QY 101 GlnAspValSerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuTyrThrCys 120
Db 346 AAAGATGCATCAATCAACATAGAAAATATGCAGTTTATACAAATGGCACCTATATCTGT 405
QY 121 AsnValSerArgGluPheGluAlaHisArgPropheValLysThrThrArgLeu 140
Db 406 GATGTCAAAAAC----- 417
QY 141 IleProLeuArgValThrGluGluAlaGlyGluAspPheThrSerValValSerGlu--- 159
Db 418 ---CCTCTGACATCGTTGCCAGCTGGACACATTAGGATCTATGTCTGTAGAAAAGAG 474
QY 160 -----IleMetMetTyrIleLeuLeuValPheLeuThr 170
Db 475 AATTTGCCGTGTTCAGTTTGGGTAGTGGTGGGCATAGTTACTGTGCTGTAGGT 534
QY 171 LeuTrpLeuLeuIleGluMetIleTyrCysTyr-----ArgLysValSerLysAla 187
Db 535 CTCACCTCTCATCAGCATGATTCGTGCTCTCTATAGAGGAAAAAACTCTAAACGG 594
QY 188 Glu 188
Db 595 GAT 597

RESULT 13

US-09-430-503-41
; Sequence 41, Application US/09430503
; Patent No. 6355786
; GENERAL INFORMATION:
; APPLICANT: Zhao, Zhizhuang
; TITLE OF INVENTION: PURIFIED AND ISOLATED PROTEIN ZERO RELATED (PZR) AND
; FILE REFERENCE: THERAPEUTIC AND SCREENING METHODS USING SAME
; CURRENT FILING DATE: 1999-10-29
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(597)
US-09-430-503-41

Alignment Scores:

| | | | |
|------------------------|----------|---------------|-----|
| Pred. No.: | 2,29e-07 | Length: | 600 |
| Score: | 127.00 | Matches: | 47 |
| Percent Similarity: | 37.57% | Conservative: | 21 |
| Best Local Similarity: | 25.97% | Mismatches: | 73 |
| Query Match: | 11.30% | Indels: | 40 |
| DB: | 3 | Gaps: | 7 |

US-09-977-579-2 (1-215) x US-09-430-503-41 (1-600)

QY 27 ValGluValProSerGluThrGluAlaValGlnGlyAsnProMetLysLeuArgCysIle 46
Db 118 GTATATACGCCAAAGAAATCTTCGTGGCAATGGTACACAAGGGAGCTGACCTGC--- 174

QY 47 SerCysMetLysArgGluGluValGluAlaThr-----ThrValValGluTrpPhe 63
Db 175 -----AAGTTCAAGTCTACTAGTACGACTGGGGGTTGACCTCAGCTCTCTCGGAGC 225
QY 64 TyrArgProGluGlyGlyLysAspPheLeu---IleTyrGluTyrArgAsnGlyHisGln 82
Db 226 TTCCAGCAGAGGGGGCCGACACTACTGTGTGCTTTTCCACTACTCCCAAGGGCAAGTG 285
QY 83 GluValGluSer-----ProPheGlnGlyArgLeuGlnTrpAsnGlySerLysAspLeu 100
Db 286 TACCTTGGGAATATCCACCAATTTAAAGACAGAGATCAGCTGGCTGGAGACCTTGACAAG 345
QY 101 GlnAspValSerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuTyrThrCys 120
Db 346 AAAGATGCATCAATCAACATAGAAAATATGCAGTTTATACAAATGGCACCTATATCTGT 405
QY 121 AsnValSerArgGluPheGluAlaHisArgPropheValLysThrThrArgLeu 140
Db 406 GATGTCAAAAAC----- 417
QY 141 IleProLeuArgValThrGluGluAlaGlyGluAspPheThrSerValValSerGlu--- 159
Db 418 ---CCTCTGACATCGTTGCCAGCTGGACACATTAGGCTCTATGTCTGTAGAAAAGAG 474
QY 160 -----IleMetMetTyrIleLeuLeuValPheLeuThr 170
Db 475 AATTTGCCGTGTTCAGTTTGGGTAGTGGTGGGCATAGTTACTGTGCTGTAGGT 534
QY 171 LeuTrpLeuLeuIleGluMetIleTyrCysTyr-----ArgLysValSerLysAla 187
Db 535 CTCACCTCTCATCAGCATGATTCGTGCTCTCTATAGAGGAAAAAACTCTAAACGG 594
QY 188 Glu 188
Db 595 GAT 597

RESULT 14

US-09-430-503-45
; Sequence 45, Application US/09430503
; Patent No. 6355786
; GENERAL INFORMATION:
; APPLICANT: Zhao, Zhizhuang
; TITLE OF INVENTION: PURIFIED AND ISOLATED PROTEIN ZERO RELATED (PZR) AND
; FILE REFERENCE: THERAPEUTIC AND SCREENING METHODS USING SAME
; CURRENT FILING DATE: 1999-10-29
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(597)
US-09-430-503-45

Alignment Scores:

| | | | |
|------------------------|----------|---------------|-----|
| Pred. No.: | 2,29e-07 | Length: | 600 |
| Score: | 127.00 | Matches: | 47 |
| Percent Similarity: | 37.57% | Conservative: | 21 |
| Best Local Similarity: | 25.97% | Mismatches: | 73 |
| Query Match: | 11.30% | Indels: | 40 |
| DB: | 3 | Gaps: | 7 |

US-09-977-579-2 (1-215) x US-09-430-503-45 (1-600)

QY 27 ValGluValProSerGluThrGluAlaValGlnGlyAsnProMetLysLeuArgCysIle 46
Db 118 GTATATACGCCAAAGAAATCTTCGTGGCAATGGTACACAAGGGAGCTGACCTGC--- 174

```

QY 47 SerCysMetLysArgGluGluValGluAlaThr-----ThrValValGluTrpPhe 63
Db 175 -----AAGTTCAAGTCTACTAGTACGACTGGCGGGTTGACCTCAGTCTCTCTGGAGC 225
QY 64 TyrArgProGluGlyGlyLysAspPheLeu---IleTyrGluTyrArgAsnGlyHisGln 82
Db 226 TCCAGCCAGAGGGGGCCGACACTACTGTGTGCTGTTTTTCCCACTACTCCCAAGGGCAAGTG 285
QY 83 GluValGluSer-----ProPheGlnGlyArgLeuGlnTrpAsnGlySerLysAspLeu 100
Db 286 TACCTGGGAATATATCCACATTTAAAGACAGAAATCAGCTGGGCTGGAGACCTTGACAAG 345
QY 101 GlnAspValSerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuTyrThrCys 120
Db 346 AAAGATGCATCAATCAACATAGAAAATATGACGTTTATACAAATGCGACCTATATCTGT 405
QY 121 AsnValSerArgGluPheGluAlaHisArgProPheValLysThrArgLeu 140
Db 406 GATGTCAAAAAC----- 417
QY 141 IleProLeuArgValThrGluGluAlaGlyGluAspPheThrSerValValSerGlu--- 159
Db 418 ---CCTCCTGACATCGTTGTCAGCCTGGACACATTAGGATCTATGTCGTAGAAAAGAG 474
QY 160 -----IleMetMetTyrIleLeuLeuValPheLeuThr 170
Db 475 AATTTGCTGTGTTTCCAGTTTGGGTAGTGGTGGCATAGTTACTGCTGTGGTCTTAGGT 534
QY 171 LeuTrpLeuLeuIleGluMetIleTyrCysTyr-----ArgLysValSerLysAla 187
Db 535 CTCACCTGCTCATCAGCATGATTCTGGCTGTCTCTATAGAGGAAAAAACTCTAAACGG 594
QY 188 Glu 188
Db 595 GAT 597

```

RESULT 15

```

US-09-430-503-17
; Sequence 17, Application US/09430503
; Patent No. 6355786
; GENERAL INFORMATION:
; APPLICANT: Zhao, Zhizhuang
; TITLE OF INVENTION: PURIFIED AND ISOLATED PROTEIN ZERO RELATED (PZR) AND
; TITLE OF INVENTION: THERAPEUTIC AND SCREENING METHODS USING SAME
; FILE REFERENCE: Attorney Docket No. 6355786 1242-11/2
; CURRENT APPLICATION NUMBER: US/09/430,503
; CURRENT FILING DATE: 1999-10-29
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 752
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(627)
US-09-430-503-17

```

Alignment Scores:

```

Pred. No.: 3.28e-07 Length: 752
Score: 127.00 Matches: 47
Percent Similarity: 37.57% Conservative: 21
Best Local Similarity: 25.97% Mismatches: 73
Query Match: 11.30% Indels: 40
DB: 3 Gaps: 7

```

US-09-977-579-2 (1-215) x US-09-430-503-17 (1-752)

```

QY 27 ValGluValProSerGluThrGluAlaValGlnGlyAsnProMetLysLeuArgCysIle 46
Db 118 GTATATACGCAAAAGAAATCTTCGTGGCAAAATGTTACACAGGGAAAGCTGACCTGC--- 174
QY 47 SerCysMetLysArgGluGluValGluAlaThr-----ThrValValGluTrpPhe 63

```

Search completed: April 7, 2005, 17:12:30
Job time : 172 secs

```

Db 175 -----AAGTTCAAGTCTACTAGTACGACTGGCGGGTTGACCTCAGTCTCTCTGGAGC 225
QY 64 TyrArgProGluGlyGlyLysAspPheLeu---IleTyrGluTyrArgAsnGlyHisGln 82
Db 226 TCCAGCCAGAGGGGGCCGACACTACTGTGTGCTGTTTTTCCCACTACTCCCAAGGGCAAGTG 285
QY 83 GluValGluSer-----ProPheGlnGlyArgLeuGlnTrpAsnGlySerLysAspLeu 100
Db 286 TACCTGGGAATATATCCACATTTAAAGACAGAAATCAGCTGGGCTGGAGACCTTGACAAG 345
QY 101 GlnAspValSerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuTyrThrCys 120
Db 346 AAAGATGCATCAATCAACATAGAAAATATGACGTTTATACAAATGCGACCTATATCTGT 405
QY 121 AsnValSerArgGluPheGluAlaHisArgProPheValLysThrArgLeu 140
Db 406 GATGTCAAAAAC----- 417
QY 141 IleProLeuArgValThrGluGluAlaGlyGluAspPheThrSerValValSerGlu--- 159
Db 418 ---CCTCCTGACATCGTTGTCAGCCTGGACACATTAGGCTCTATGTCGTAGAAAAGAG 474
QY 160 -----IleMetMetTyrIleLeuLeuValPheLeuThr 170
Db 475 AATTTGCTGTGTTTCCAGTTTGGGTAGTGGTGGCATAGTTACTGCTGTGGTCTTAGGT 534
QY 171 LeuTrpLeuLeuIleGluMetIleTyrCysTyr-----ArgLysValSerLysAla 187
Db 535 CTCACCTGCTCATCAGCATGATTCTGGCTGTCTCTATAGAGGAAAAAACTCTAAACGG 594
QY 188 Glu 188
Db 595 GAT 597

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: April 7, 2005, 15:19:45 ; Search time 502 Seconds
(without alignments)
2595.420 Million cell updates/sec

Title: US-09-977-579-2
Perfect score: 1124
Sequence: 1 MPANRFLPLASLVLIYWS.....SDYLAIPSENKENSAPVVEE 215

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0
Searched: 5615251 seqs, 3030001701 residues

Total number of hits satisfying chosen parameters: 11230502

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USFIO_spool_p/US09977579/runat_06042005_163458_23652/app_query.fasta_1.391
-DB=Published Applications NA -OFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALL=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09977579 @CGN 1 1 723 @runat 06042005 163458 23652
-NCPU=6 -ICPU=3 -NO MAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 21: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 22: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---------------------|
| 1 | 1124 | 100.0 | 1261 | 11 | US-09-977-579-4 |
| 2 | 1105 | 98.3 | 645 | 13 | US-10-029-191-21 |
| 3 | 1105 | 98.3 | 2220 | 11 | US-09-977-579-3 |
| 4 | 1105 | 98.3 | 2632 | 13 | US-10-029-191-22 |
| 5 | 1105 | 98.3 | 3108 | 13 | US-10-029-191-1 |
| 6 | 477 | 42.4 | 657 | 13 | US-10-029-191-23 |
| 7 | 473 | 42.1 | 1335 | 18 | US-10-723-860-2247 |
| 8 | 473 | 42.1 | 1414 | 18 | US-10-477-272-1 |
| 9 | 472 | 42.0 | 1490 | 9 | US-09-917-800A-1654 |
| 10 | 260.5 | 23.2 | 974 | 17 | US-10-401-916-13 |
| 11 | 258.5 | 23.0 | 807 | 17 | US-10-401-916-12 |
| 12 | 234 | 20.8 | 407 | 17 | US-10-276-774-718 |
| 13 | 167 | 14.9 | 1029 | 17 | US-10-205-331-111 |
| 14 | 165.5 | 14.7 | 3583 | 18 | US-10-723-860-6471 |
| 15 | 163.5 | 14.5 | 1929 | 10 | US-09-960-706-472 |
| 16 | 160.5 | 14.3 | 970 | 19 | US-10-874-706-84 |
| 17 | 160.5 | 14.3 | 983 | 13 | US-10-053-107-11 |
| 18 | 160.5 | 14.3 | 983 | 14 | US-10-227-884-237 |
| 19 | 160.5 | 14.3 | 983 | 14 | US-10-230-163-237 |
| 20 | 160.5 | 14.3 | 983 | 14 | US-10-230-338-237 |
| 21 | 160.5 | 14.3 | 983 | 14 | US-10-218-631-237 |
| 22 | 160.5 | 14.3 | 983 | 14 | US-10-230-414-237 |
| 23 | 160.5 | 14.3 | 983 | 14 | US-10-213-145-11 |
| 24 | 160.5 | 14.3 | 983 | 14 | US-10-232-224-237 |
| 25 | 160.5 | 14.3 | 983 | 14 | US-10-216-159A-237 |
| 26 | 160.5 | 14.3 | 983 | 14 | US-10-218-849-237 |
| 27 | 160.5 | 14.3 | 983 | 14 | US-10-227-873-237 |
| 28 | 160.5 | 14.3 | 983 | 14 | US-10-227-883-237 |
| 29 | 160.5 | 14.3 | 983 | 14 | US-10-219-076-237 |
| 30 | 160.5 | 14.3 | 983 | 14 | US-10-230-434-237 |
| 31 | 160.5 | 14.3 | 983 | 14 | US-10-213-199-11 |
| 32 | 160.5 | 14.3 | 983 | 14 | US-10-219-003-237 |
| 33 | 160.5 | 14.3 | 983 | 14 | US-10-219-075-237 |
| 34 | 160.5 | 14.3 | 983 | 14 | US-10-219-464-237 |
| 35 | 160.5 | 14.3 | 983 | 14 | US-10-219-466-237 |
| 36 | 160.5 | 14.3 | 983 | 14 | US-10-219-479-237 |
| 37 | 160.5 | 14.3 | 983 | 14 | US-10-219-481-237 |
| 38 | 160.5 | 14.3 | 983 | 14 | US-10-230-260-237 |
| 39 | 160.5 | 14.3 | 983 | 14 | US-10-232-231-237 |
| 40 | 160.5 | 14.3 | 983 | 14 | US-10-232-233-237 |
| 41 | 160.5 | 14.3 | 983 | 14 | US-10-216-165-237 |
| 42 | 160.5 | 14.3 | 983 | 14 | US-10-218-956-237 |
| 43 | 160.5 | 14.3 | 983 | 14 | US-10-219-468-237 |
| 44 | 160.5 | 14.3 | 983 | 14 | US-10-219-478-237 |
| 45 | 160.5 | 14.3 | 983 | 14 | US-10-219-536-237 |

ALIGNMENTS

RESULT 1

US-09-977-579-4
; Sequence 4, Application US/09977579
; Publication No. US20040248240A1
; GENERAL INFORMATION:
; APPLICANT: Cambridge University Technical Services
; TITLE OF INVENTION: A novel family of beta sub-unit proteins from a voltage gated so
; TITLE OF INVENTION: channel
; FILE OF INVENTION: nucleic acids encoding them and therapeutic or diagnostic uses t
; FILE REFERENCE: 674558-2001
; CURRENT APPLICATION NUMBER: US/09/977,579
; PRIOR FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: PCT/EP00/01783
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60,129,473
; PRIOR FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1261

; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-977-579-4

Alignment Scores:
Pred. No.: 2,02e-143 Length: 1261
Score: 1124.00 Matches: 215
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 11 Gaps: 0

US-09-977-579-2 (1-215) x US-09-977-579-4 (1-1261)

```
QY 1 MetProAlaPheAsnArgLeuPheProLeuAlaSerLeuValLeuIleTyrTrpValSer 20
DB 376 ATGCCTGCTCAATAGATTGTTCCCTCGGCTTCTCTCGTCTTATCTAGTGGTCACT 435

QY 21 ValCysPheProValCysValGluValProSerGluThrGluAlaValGlnGlyAsnPro 40
DB 436 GTCTGCTCCCTGTGTGTGGAGTGCCTCGGAGACGGAGCGCGTGAGGGCAACCC 495

QY 41 MetLysLeuArgCysIleSerCysMetLysArgGluGluValGluAlaThrThrValVal 60
DB 496 ATGAAGCTCGCTGCATCTCTGTCATGAAGAGAGAGAGGTGGAGGCCACCAACGGTGGTG 555

QY 61 GluTrpPheTyrArgProGluGlyGlyLysAspPheLeuIleTyrGluTyrArgAsnGly 80
DB 556 GAATGGTCTTACAGGCCCGAGGGCGGTAAGATTTCCTTTATTACGAGTATCGGAATGCG 615

QY 81 HisGlnGluValGluSerProPheGlnGlyArgLeuGlnTrpAsnGlySerLysAspLeu 100
DB 616 CACCAGGAGGTGGAGAGCCCTTTACAGGGCGCCTGCAGTGAATGGCAGGACACCTG 675

QY 101 GlnAspValSerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuTyrThrCys 120
DB 676 CAGGAGCTGTCATCTGTCACGTCATCTGNAAGCATCTGCGCCTCTACACCTGC 735

QY 121 AsnValSerArgGluPheGluPheGluAlaHisArgProPheValLysThrThrArgLeu 140
DB 736 AATGTGTCGGGAGTTTGAGTTGAGGGCGCATCGGCCCTTTGTGAAGACGACGGCGCTG 795

QY 141 IleProLeuArgValThrGluGluAlaGlyGluAspPheThrSerValValSerGluIle 160
DB 796 ATCCCCCTAAGAGTCAACGAGAGGCTGGAGAGGACTTCACTCTGTGTGCTCAGAAATC 855

QY 161 MetMetTyrIleLeuLeuValPheLeuThrLeuTrpLeuLeuIleGluMetIleTyrCys 180
DB 856 ATGATGTATACATCTTCTGTCTTCTCACCCTGTGGCTGCTCATCGAGATGATATATTGC 915

QY 181 TyrArgLysValSerLysAlaGluGluAlaGlnGluAsnAlaSerAspTyrLeuAla 200
DB 916 TACAGAAAGGTCTCAAAAGCCGAAGAGGAGCCCAAGAAACCGCTCTCACTACCTTGGC 975

QY 201 IleProSerGluAsnLysGluAsnSerAlaValProValGluGlu 215
DB 976 ATCCCATCTGAGAACCAAGGAGAACTCTGCGTACCGAGTGGAGAA 1020
```

RESULT 2

US-10-029-191-21
; Sequence 21, Application US/10029191
; Publication No. US20020160453A1
; GENERAL INFORMATION:
; APPLICANT: CURTIS, Rory A. J.
; TITLE OF INVENTION: NOVEL GENE ENCODING A SODIUM CHANNEL BETA-3 SUBUNIT
; FILE REFERENCE: 210147.00XX/5U1
; CURRENT APPLICATION NUMBER: US/10/029,191
; CURRENT FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 09/569,978
; PRIOR FILING DATE: 2000-05-12
; PRIOR APPLICATION NUMBER: US 60/134,198
; PRIOR FILING DATE: 1999-05-14

RESULT 3

US-09-977-579-3
; Sequence 3, Application US/09977579
; Publication No. US20040248240A1
; GENERAL INFORMATION:
; APPLICANT: Cambridge University Technical Services
; TITLE OF INVENTION: A novel family of beta sub-unit proteins from a voltage gated sodi
; TITLE OF INVENTION: channel
; TITLE OF INVENTION: nucleic acids encoding them and therapeutic or diagnostic uses th
; FILE REFERENCE: 674558-2001
; CURRENT APPLICATION NUMBER: US/09/977,579

; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 645
; TYPE: DNA
; ORGANISM: Rattus sp.
US-10-029-191-21

Alignment Scores:

Pred. No.: 2,91e-141 Length: 645
Score: 1105.00 Matches: 211
Percent Similarity: 98.14% Conservative: 0
Best Local Similarity: 98.14% Mismatches: 4
Query Match: 98.31% Indels: 0
DB: 13 Gaps: 0

US-09-977-579-2 (1-215) x US-10-029-191-21 (1-645)

```
QY 1 MetProAlaPheAsnArgLeuPheProLeuAlaSerLeuValLeuIleTyrTrpValSer 20
DB 1 ATGCCTGCGCTTCAACAGATTGCTTCCCTTAGCTTCTCTAGTGTCTCATCTACTGGGTGCGA 60

QY 21 ValCysPheProValCysValGluValProSerGluThrGluAlaValGlnGlyAsnPro 40
DB 61 GTCTGCTTCCCTGTGTGTGGAGTGCCTCGGAGACAGAGCGGTGCGGGCAATCC 120

QY 41 MetLysLeuArgCysIleSerCysMetLysArgGluGluValGluAlaThrThrValVal 60
DB 121 ATGAAGCTGAGTGCATCTCTGTCATGAAGAGGAGGAGGTGGAGGCCACCACTGTGGTG 180

QY 61 GluTrpPheTyrArgProGluGlyGlyLysAspPheLeuIleTyrGluTyrArgAsnGly 80
DB 181 GAGTGGTCTACAGGCCCTGAGGGCGGTAAGATTTCCTTATATATAGTATCGGAATGCG 240

QY 81 HisGlnGluValGluSerProPheGlnGlyArgLeuGlnTrpAsnGlySerLysAspLeu 100
DB 241 CACCAGGAGGTGGAGAGCCCTTCCAGGCCGTCTGCAGTGAATGGGAGCAAGACCTG 300

QY 101 GlnAspValSerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuTyrThrCys 120
DB 301 CAGGAGCTATCCATCACTGTCATCAATGTCATCTTGAATGACTCTGGCCTCTACACATGC 360

QY 121 AsnValSerArgGluPheGluPheGluAlaHisArgProPheValLysThrThrArgLeu 140
DB 361 AATGTGTCCAGGAGTTCGAATTCGAGGCACACAGGCCCTTTGTGAAGACCAACGAGACTG 420

QY 141 IleProLeuArgValThrGluGluAlaGlyGluAspPheThrSerValValSerGluIle 160
DB 421 ATACCTTTGCGAGTCACTGAAGAGCGGGAGAGACTTCACCTCCGTGGTCTCGGAAATC 480

QY 161 MetMetTyrIleLeuLeuValPheLeuThrLeuTrpLeuLeuIleGluMetIleTyrCys 180
DB 481 ATGATGTATACCTCTCTGTCTTCTCCTCACTTGTGGCTGTTTATTGAGATGATCATTTGC 540

QY 181 TyrArgLysValSerLysAlaGluGluAlaGlnGluAsnAlaSerAspTyrLeuAla 200
DB 541 TACAGAAAGGTCTCAAGGCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600

QY 201 IleProSerGluAsnLysGluAsnSerAlaValProValGluGlu 215
DB 601 ATCCCTTTCAGAGAACAGAGGAGAACTCTGTGTGTACCTGTGGAGGAA 645
```

APPLICANT: Cambridge University Technical Services

; TITLE OF INVENTION: A novel family of beta sub-unit proteins from a voltage gated sodi
; TITLE OF INVENTION: channel
; TITLE OF INVENTION: nucleic acids encoding them and therapeutic or diagnostic uses th
; FILE REFERENCE: 674558-2001
; CURRENT APPLICATION NUMBER: US/09/977,579

; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: PCT/EP00/01783
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60,129,473
; PRIOR FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 2220
; TYPE: DNA
; ORGANISM: rat
US-09-977-579-3

Alignment Scores:
Pred. No.: 1.98e-140 Length: 2220
Score: 1105.00 Matches: 211
Percent Similarity: 98.14% Conservative: 0
Best Local Similarity: 98.14% Mismatches: 4
Query Match: 98.31% Indels: 0
DB: 11 Gaps: 0

US-09-977-579-2 (1-215) x US-09-977-579-3 (1-2220)

| | | | |
|----|-----|---|------|
| QY | 1 | MetProAlaPheAsnArgLeuPheProLeuAlaSerLeuValLeuIleTyrTrpValSer | 20 |
| DB | 363 | ATGCTGCTCCCTTCAACAGATTGCTTCCCTAGCTTCTCTAGTGCTCATCTACTGGGTGAGA | 422 |
| QY | 21 | ValCysPheProValCysValGluValProSerGluThrGluAlaValGlnGlyAsnPro | 40 |
| DB | 423 | GTCTGCTTCCCTGTGTGTGGAGTGCCTCCGAGACAGAGCGGTGAGGGCAATCCC | 482 |
| QY | 41 | MetLysLeuArgCysIleSerCysMetLysArgGluValGluAlaThrValVal | 60 |
| DB | 483 | ATGAGCTGAGGTGCATCTCTGCATGAGAGGGAGGAGGTGAGGCCACCATCTGGTG | 542 |
| QY | 61 | GluTrpPheTyrArgProGluGlyGlyAspPheLeuIleTyrGluTyrArgAsnGly | 80 |
| DB | 543 | GAGTGTGTTCTACAGGCTGAGGCGGTAAAGATTCTCTATATATGAGTATCGGAATGGC | 602 |
| QY | 81 | HisGlnGluValGluSerProPheGlnGlyArgLeuGlnTrpAsnGlySerLysAspLeu | 100 |
| DB | 603 | CACAGAGAGTGAGAGACCCCTTCCAGGCGGTCTGCAGTGGAAATGGGAGCAAGACCTG | 662 |
| QY | 101 | GlnAspValSerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuTyrThrCys | 120 |
| DB | 663 | CAGACGTATCCATCACTGACTCAATGTCAATGTCAATGTCAATGTCAATGTCAATGC | 722 |
| QY | 121 | AsnValSerArgGluPheGluPheGluAlaHisArgProPheValLysThrArgLeu | 140 |
| DB | 723 | AATGTGTCCAGGAGTTCGAATTCGAGGCACACAGGCTTTTGTGAAGACCCAGAGACTG | 782 |
| QY | 141 | IleProLeuArgValThrGluGluAlaGlyGluAspPheThrSerValValSerGluIle | 160 |
| DB | 783 | ATACCTTTGGAGTCACTGAAGAGCGGGAGAGACTTCACTCCGCTGGGTCTCGGAATC | 842 |
| QY | 161 | MetMetTyrIleLeuLeuValPheLeuThrLeuTrpLeuLeuIleGluMetIleTyrCys | 180 |
| DB | 843 | ATGATGTACATCTCTCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT | 902 |
| QY | 181 | TyrArgLysValSerLysAlaGluGluAlaGlnGluAsnAlaSerAspTyrLeuAla | 200 |
| DB | 903 | TACAGAAAGGTCTCTAAGCGCGAAGAGGAGCAGAGGAAATGCGTCTGACTACCTTGCT | 962 |
| QY | 201 | IleProSerGluAsnLysGluAsnSerAlaValProValGluGlu | 215 |
| DB | 963 | ATCCCTTTCAGAGAACAGGAGAACTCTGTGGTACCTGTGGAGAA | 1007 |

RESULT 4

US-10-029-191-22

; Sequence 22, Application US/10029191

; Publication No. US20020160453A1

; GENERAL INFORMATION:

; APPLICANT: CURTIS, RORY A.J.

; TITLE OF INVENTION: NOVEL GENE ENCODING A SODIUM CHANNEL BETA-3 SUBUNIT
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 210147.00XX/5U1
; CURRENT APPLICATION NUMBER: US/10/029,191
; CURRENT FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 09/569,978
; PRIOR FILING DATE: 2000-05-12
; PRIOR APPLICATION NUMBER: US 60/134,198
; PRIOR FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 2632
; TYPE: DNA
; ORGANISM: Rattus sp.
US-10-029-191-22

Alignment Scores:
Pred. No.: 2.57e-140 Length: 2632
Score: 1105.00 Matches: 211
Percent Similarity: 98.14% Conservative: 0
Best Local Similarity: 98.14% Mismatches: 4
Query Match: 98.31% Indels: 0
DB: 13 Gaps: 0

US-09-977-579-2 (1-215) x US-10-029-191-22 (1-2632)

| | | | |
|----|-----|---|-----|
| QY | 1 | MetProAlaPheAsnArgLeuPheProLeuAlaSerLeuValLeuIleTyrTrpValSer | 20 |
| DB | 78 | ATGCTGCTCCCTTCAACAGATTGCTTCCCTAGCTTCTCTAGTGCTCATCTACTGGGTGAGA | 137 |
| QY | 21 | ValCysPheProValCysValGluValProSerGluThrGluAlaValGlnGlyAsnPro | 40 |
| DB | 138 | GTCTGCTTCCCTGTGTGTGGAGTGCCTCCGAGACAGAGCGGTGAGGGCAATCCC | 197 |
| QY | 41 | MetLysLeuArgCysIleSerCysMetLysArgGluValGluAlaThrValVal | 60 |
| DB | 198 | ATGAGCTGAGGTGCATCTCTGCATGAGAGGGAGGAGGTGAGGCCACCATCTGGTG | 257 |
| QY | 61 | GluTrpPheTyrArgProGluGlyGlyAspPheLeuIleTyrGluTyrArgAsnGly | 80 |
| DB | 258 | GAGTGTGTTCTACAGGCTGAGGCGGTAAAGATTCTCTATATGAGTATCGGAATGGC | 317 |
| QY | 81 | HisGlnGluValGluSerProPheGlnGlyArgLeuGlnTrpAsnGlySerLysAspLeu | 100 |
| DB | 318 | CACAGAGAGTGAGAGACCCCTTCCAGGCGGTCTGCAGTGGAAATGGGAGCAAGACCTG | 377 |
| QY | 101 | GlnAspValSerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuTyrThrCys | 120 |
| DB | 378 | CAGACGTATCCATCACTGACTCAATGTCAATGTCAATGTCAATGTCAATGTCAATGC | 437 |
| QY | 121 | AsnValSerArgGluPheGluPheGluAlaHisArgProPheValLysThrArgLeu | 140 |
| DB | 438 | AATGTGTCCAGGAGTTCGAATTCGAGGCACACAGGCTTTTGTGAAGACCCAGAGACTG | 497 |
| QY | 141 | IleProLeuArgValThrGluGluAlaGlyGluAspPheThrSerValValSerGluIle | 160 |
| DB | 498 | ATACCTTTGGAGTCACTGAAGAGCGGGAGAGACTTCACTCCGCTGGGTCTCGGAATC | 557 |
| QY | 161 | MetMetTyrIleLeuLeuValPheLeuThrLeuTrpLeuLeuIleGluMetIleTyrCys | 180 |
| DB | 558 | ATGATGTACATCTCTCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT | 617 |
| QY | 181 | TyrArgLysValSerLysAlaGluGluAlaGlnGluAsnAlaSerAspTyrLeuAla | 200 |
| DB | 618 | TACAGAAAGGTCTCTAAGCGCGAAGAGGAGCAGAGGAAATGCGTCTGACTACCTTGCT | 677 |
| QY | 201 | IleProSerGluAsnLysGluAsnSerAlaValProValGluGlu | 215 |
| DB | 678 | ATCCCTTTCAGAGAACAGGAGAACTCTGTGGTACCTGTGGAGAA | 722 |

RESULT 5

US-10-029-191-1

Sequence 1, Application US/10029191
Publication No. US20020160453A1

GENERAL INFORMATION:

APPLICANT: CURTIS, RORY A.J.

TITLE OF INVENTION: NOVEL GENE ENCODING A SODIUM CHANNEL BETA-3 SUBUNIT

TITLE OF INVENTION: PROTEIN

FILE REFERENCE: 210147.00XX/5U1

CURRENT APPLICATION NUMBER: US/10/029,191

CURRENT FILING DATE: 2001-12-20

PRIOR APPLICATION NUMBER: 09/569,978

PRIOR FILING DATE: 2000-05-12

PRIOR APPLICATION NUMBER: US 60/134,198

PRIOR FILING DATE: 1999-05-14

NUMBER OF SEQ ID NOS: 23

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 1

LENGTH: 3108

TYPE: DNA

ORGANISM: Rattus sp.

US-10-029-191-1

Alignment Scores:

Pred. No.: 3,32e-140 Length: 3108
Score: 1105.00 Matches: 211
Percent Similarity: 98.14% Conservatives: 0
Best Local Similarity: 98.14% Mismatches: 4
Query Match: 98.31% Indels: 0
DB: 13 Gaps: 0

US-09-977-579-2 (1-215) x US-10-029-191-1 (1-3108)

```
QY 1 MetProAlaPheAsnArgLeuPheProLeuAlaSerLeuValLeuLeuLeuTyrTrpValSer 20
DB 78 ATGCCTGCTTCAACAGATGCTTCCCTAGCTTCTTAGTGCTCATCTACTGGGTGCGA 137
QY 21 ValCysPheProValCysValGluValProSerGluThrGluAlaValGlnGlyAsnPro 40
DB 138 GTCTGCTTCCCTGTGTGTGAAGTGCCTCGAGACAGACGGGTGAGGGCAATGCC 197
QY 41 MetLysLeuArgCysIleSerCysMetLysArgGluGluValGluAlaThrThrValVal 60
DB 198 ATGAAGCTCAGGTGCATCTCTGTCATGAAGAGGAGGAGTGGAGGCCACCACTGTGGTG 257
QY 61 GluTrpPheTyrArgProGluGlyLysAspPheLeuLeuLeuTyrGluTyrArgAsnGly 80
DB 258 GAGTGGTCTTACAGGCTCAGGCGGTAAAGATTTCTTTATATATAGTATCGGAATGCC 317
QY 81 HisGlnGluValGluSerProPheGlnGlyArgLeuGlnTrpAsnGlySerLysAspLeu 100
DB 318 CACCAGGAAGTGGAGAGCCCTTCCAGGCCGTCTGCACTGGATGGAGGCAAGACCTG 377
QY 101 GinAspValSerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuTyrThrCys 120
DB 378 CAGGAGTATCATCACTACTCAATGTCACTTTGAATGACTCTGGCCCTCTACACATGC 437
QY 121 AsnValSerArgGluPheGluPheGluAlaHisArgProPheValLysThrThrArgLeu 140
DB 438 AATGTGTCCAGGAGTTCGAATTCGAGGCGACACAGGCCCTTTTGTGAAGACCAACGACTG 497
QY 141 IleProLeuArgValThrGluAlaGlyGluAspPheThrSerValValSerGluIle 160
DB 498 ATACCTTTCCGAGTCACTGAAGAGCGGAGAGAGACTTCACCTCCGTGTCTCGGAATC 557
QY 161 MetMetTyrIleLeuLeuValPheLeuThrLeuTyrProLeuLeuIleGluMetIleTyrCys 180
DB 558 ATGATGTACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 617
QY 181 TyrArgLysValSerLysAlaGluAlaAlaGlnGluAsnAlaSerAspTyrIleuAla 200
DB 618 TACAGAAAGGTCTCTAAGGCCGGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 677
QY 201 IleProSerGluAsnLysGluAsnSerAlaValProValGluGlu 215
```

Db 678 ATCCCTTCAGACACAGGAGAACTCTGTGTACTCTGTGGAGAA 722

RESULT 6

US-10-029-191-23

Sequence 23, Application US/10029191

Publication No. US20020160453A1

GENERAL INFORMATION:

APPLICANT: CURTIS, RORY A.J.

TITLE OF INVENTION: NOVEL GENE ENCODING A SODIUM CHANNEL BETA-3 SUBUNIT

TITLE OF INVENTION: PROTEIN

FILE REFERENCE: 210147.00XX/5U1

CURRENT APPLICATION NUMBER: US/10/029,191

CURRENT FILING DATE: 2001-12-20

PRIOR APPLICATION NUMBER: 09/569,978

PRIOR FILING DATE: 2000-05-12

PRIOR APPLICATION NUMBER: US 60/134,198

PRIOR FILING DATE: 1999-05-14

NUMBER OF SEQ ID NOS: 23

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 23

LENGTH: 657

TYPE: DNA

ORGANISM: Lepus Sp.

US-10-029-191-23

Alignment Scores:

Pred. No.: 5.43e-55 Length: 657
Score: 477.00 Matches: 106
Percent Similarity: 64.49% Conservatives: 32
Best Local Similarity: 49.53% Mismatches: 68
Query Match: 42.44% Indels: 8
DB: 13 Gaps: 5

US-09-977-579-2 (1-215) x US-10-029-191-23 (1-657)

```
QY 10 LeuAlaSerLeuValLeuLeuTyrTrpValSerValCysPheProValCysValGluVal 29
DB 13 CTGGCTTTCGTGGTGGCGCCGCTGTGCTCTCGGCTGGGGGGTCTGCTGGAGGTG 72
QY 30 ProSerGluThrGluAlaValGlnGlyAsnProMetLysLeuArgCysIleSerCysMet 49
DB 73 GACTCGGAGACGAGCGGTGTACGGGATGACCTTCAAAATTCGTGCACTCTCTGCAAG 132
QY 50 LysArgGluGluValGluAlaThrThrValValGluTyrPheTyrArgProGluGlyGly 69
DB 133 CCGCCGACGAGACACCGCCGAGACCTTCCGAGGTGGACCTTCCGCCAGAGGCACT 192
QY 70 LysAspPheLeu---IleTyrGluTyrArgAsnGlyHisGlnGluValGluSerPro--- 87
DB 193 GAGGAGTTCGTCAAGATCTCGCTATGAGACGAGGTGCTGCTGAGGAGACGAG 252
QY 88 ---PheGlnGlyArgLeuGlnTrpAsnGlySer-----LysAspLeuGlnAspVal 103
DB 253 CGCTTTGAGGGCGCGTGTGTGGAACGCGACCGCGGGGACCAAGACCTGCGAGACCTG 312
QY 104 SerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuTyrThrCysAsnValSer 123
DB 313 TCCATCTTCATCACCATGTTCACCTACACCACTCGGGGACTACCACTGCGCATGTCTAC 372
QY 124 ArgGluPheGluPheGluAlaHisArgProPheValLysThrThrArgLeuIleProLeu 143
DB 373 CGCTGCTCTCTCTCGAATACTACGAGCACACACCGGCTGCTCAAGAGATCCACCTG 432
QY 144 ArgValThrGluGluAlaGlyGluAspPheThrSerValSerGluIleMetMetTyr 163
DB 433 GAGGTGGTGGACAAAGCCCAACAGACATGCGATCCATCGTGTGCGAGATCATGATGTAC 492
QY 164 IleLeuLeuValPheLeuThrLeuTyrPheLeuLeuIleGluMetIleTyrCysTyrArgLys 183
DB 493 GTGCTCATCTGTGTGTGACCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 552
QY 184 ValSerLysAla---GluGluAlaAlaGlnGluAsnAlaSerAspTyrIleuAlaIlePro 202
```


QY 157 ValSerGluLeuMetTyrLeuLeuValPheLeuThrLeuTrpLeuLeuGlu 176
DB 572 GTGCTCAGATCATGATGATGCTCAATGGTGTGACCATGCTCGTGGCAGAG 631
QY 177 MetIleTyrCysTyrArgLysValSerLysAla---GluGluAlaGlnGluAla 195
DB 632 ATGATTTACTGTACAGAGATGCTGCGGCCACGAGACTGCTGCACAGAGATGCC 691
QY 196 SerAspTyrLeuAlaIleProSerGluAsnLysGluAsn---SerAlaValProValGlu 214
DB 692 TCGGAATACCTGGCCATCACTCTGAAGACCAAGAGAACTGCACGGGGCTCCAGGTGGCC 751
QY 215 Glu 215
DB 752 GAA 754

RESULT 9

US-09-917-800A-1654
; Sequence 1654, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5038-US
; CURRENT APPLICATION NUMBER: US/09/917,800A
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,645
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/292,336
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1654
; LENGTH: 1490
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM_017288
US-09-917-800A-1654

Alignment Scores:
Pred. No.: 9,38e-54 Length: 1490
Score: 472.00 Matches: 105
Percent Similarity: 64.49% Conservative: 33
Best Local Similarity: 49.07% Mismatches: 68
Query Match: 41.99% Indels: 8
DB: 9 Gaps: 5

US-09-977-579-2 (1-215) x US-09-917-800A-1654 (1-1490)

QY 10 LeuAlaSerLeuValLeuIleTyrTrpValSerValCysPheProValCysValGluVal 29
DB 232 CTGGCTCTCGTGGGGCGGTGCTGTATCTCAGCCTGGGGGGGCTCGCTGGAGGTG 291

QY 30 ProSerGluThrGluAlaValGlnGlnGlyAsnProMetLysLeuArgCysAlaSerCysMet 49
DB 292 GATTCGTGACCGAGCGAGTGTATGGATGACCTTTCAAAATCCTGTGTATCTCTCTAAG 351
QY 50 LysArgGluGluValGluAlaThrThrValValGluTyrPheTyrArgProGluGlyGly 69
DB 352 COTCTGTAGTGAGACCAACCGCCGAGACCTTCACGGAGTGGACCTTCGCCAGAAAGGGCACA 411
QY 70 LysAspPheLeu---IleTyrGluTyrArgAsnGlyHisGlnGluValGluSerPro--- 87
DB 412 GAGGAATTTGTCAAGATCCTACGCTATGAGAATGAGTGTCTCAGCTGGAGGAAGATGAG 471
QY 88 ---PheGlnGlyArgLeuGlnTrpAsnGlySer-----LysAspLeuGlnAspVal 103
DB 472 CGCTTTTGAAGGGCGGTGTGGTGTGGAACGGTAGTTCGGGGCCACCAAGGACCTGCAGGACCTG 531
QY 104 SerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuTyrThrCysAsnValSer 123
DB 532 TCCATCTTCATCACCATGTGTACCTACACCACTCTGGGAGTACGAATGTCCAGTCTAC 591
QY 124 ArgGluPheGluPheGluAlaHisArgProPheValLysThrArgLeuIleProLeu 143
DB 592 CGTCTCTCTCTTCTTTGATAATTACGAGCACACACCGCTCGTCAAGAAGATCCACCTG 651
QY 144 ArgValThrGluGluAlaGlyGluAspPheThrSerValValSerGluIleMetTyr 163
DB 652 GAGGTGGTGGCAAGGCCAACAGAGATATGGCATCCATCGTGTGAGAGATCATGTATGAT 711
QY 164 IleLeuLeuValPheLeuThrLeuTrpLeuIleGluMetIleTyrCysTyrArgLys 183
DB 712 GTGCTCATTTGGTGTGTTAACCATATGGCTCGTGGCGGAGATGGTGTACTGCTACAGAAG 771
QY 184 ValSerLysAla---GluGluAlaGlnGlnGlyAsnAlaSerAspTyrLeuAlaIlePro 202
DB 772 ATTGTGCTGCCACGGAAGTGTGTCACAAGAGAATGCTCGGAATACCTGGCCATTACT 831
QY 203 SerGluAsnLysGluAsn---SerAlaValProValGluGlu 215
DB 832 TCCGAGACCAAGAGAACTGTACAGGCGTCCAGGTGGGTGAA 873

RESULT 10

US-10-401-916-13
; Sequence 13, Application US/10401916
; Publication No. US20040002439A1
; GENERAL INFORMATION:
; APPLICANT: Qin, Ning
; APPLICANT: Codd, Ellen
; APPLICANT: D'Andrea, Michael
; TITLE OF INVENTION: DNAs encoding human betala sodium channel subunit
; FILE REFERENCE: ORT-1221
; CURRENT APPLICATION NUMBER: US/10/401,916
; CURRENT FILING DATE: 2003-03-28
; PRIOR APPLICATION NUMBER: US/09/875,456A
; PRIOR FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 13
; LENGTH: 974
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-401-916-13

Alignment Scores:
Pred. No.: 5.45e-25 Length: 974
Score: 260.50 Matches: 67
Percent Similarity: 56.13% Conservative: 20
Best Local Similarity: 43.23% Mismatches: 59
Query Match: 23.18% Indels: 9
DB: 17 Gaps: 5

US-09-977-579-2 (1-215) x US-10-401-916-13 (1-974)

Qy 3 AlaPheAsnArgLeuPheProLeuAlaSerLeuValLeuLeuIleTyrTrpValSerValCys 22
 Db 1 GCCATGGGAGGCTG-----CTGGCTTAGTGTGGCGCGGCACTGCTCTCAGCC 54
 Qy 23 PheProValCysValGluValProSerGluThrGluAlaValGlnGlyAsnProMetLys 42
 Db 55 TGCAGGGGCTCGCTGGAGGTGGACTCGGAGACCGAGGCGCGTGTATGGGATGACCTTCAAA 114
 Qy 43 LeuArgCysIleSerCysMetLysArgGluValGluAlaThrThrValValGluTrp 62
 Db 115 ATTCTTTGATCTCTCTGAGCGCGGAGGAGACCAACGCTGAGACCTTACCGAGTGG 174
 Qy 63 PheTyrArgProGluGlyLysAspPheLeu---IleTyrGluTyrArgAsnGlyHis 81
 Db 175 ACCTTCGCCAGAGGCACTGAGAGTTTGTCAAGATCTCGCGCTATGAGATGAGGTG 234
 Qy 82 GlnGluValGluSerPro-----PheGlnGlyArgLeuGlnTrpAsnGlySer----- 97
 Db 235 TTGCAGCTGGAGGAGGATGAGCGCTTCGAGGCGCGCTGTGGATGGCAGCGCGGCG 294
 Qy 98 ---LysAspLeuGlnAspValSerIleThrValLeuAsnValThrLeuAsnAspSerGly 116
 Db 295 ACCAAGACCTTCAGGATCTGTCTATCTTATCACCACCAATGTACCTACACCACTCGGCG 354
 Qy 117 LeuTyrThrCysAsnValSerArgGluPheGluAlaHisArgProPheValLys 136
 Db 355 GACTACGAGTGCCAGCTACCGCGCTGCTCTTCTCGAAAACCTACGAGCACACACCGAG 414
 Qy 137 ThrThrArgLeuIleProLeuArgValThrGluAlaGlyGlu 151
 Db 415 GTCGTCAAGAAGATCCACATTTAGGTAGTGACAAA---GGTGAG 456

RESULT 11

US-10-401-916-12
 ; Sequence 12, Application US/10401916
 ; Publication No. US20040002439A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Qin, Ning
 ; APPLICANT: Codd, Ellen
 ; APPLICANT: D'Andrea, Michael
 ; TITLE OF INVENTION: DNAs encoding human betala sodium channel subunit
 ; FILE REFERENCE: ORT-1221
 ; CURRENT APPLICATION NUMBER: US/10/401.916
 ; PRIOR FILING DATE: 2003-03-28
 ; PRIOR APPLICATION NUMBER: US/09/875,456A
 ; PRIOR FILING DATE: 2001-09-10
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 12
 ; LENGTH: 807
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-401-916-12

Alignment Scores:
 Pred. No.: 7,67e-25 Length: 807
 Score: 258.50 Matches: 64
 Percent Similarity: 56.76% Conservative: 20
 Best Local Similarity: 43.24% Mismatches: 57
 Query Match: 23.00% Indels: 7
 DB: 17 Gaps: 4

US-09-977-579-2 (1-215) x US-10-401-916-12 (1-807)

Qy 10 LeuAlaSerLeuValLeuIleTyrTrpValSerValCysPheProValCysValGluVal 29
 Db 13 CTGGCTTAGTGTGGCGGCGGCACTGCTGTCTCTCAGCGCGGCGCTCGTGGAGGTG 72
 Qy 30 ProSerGluThrGluAlaValGlnGlyAsnProMetLysLeuArgCysIleSerCysMet 49
 Db 73 GACTCGGAGACCGGCGGTATGGATGACCTTCAAAATCTTTGCATCTCTCTGCAAG 132
 Qy 50 LysArgGluValGluAlaThrThrValValGluTrpPheTyrArgProGluGlyGly 69

Db 133 CGCGCAGCGAGACCAACGCTGAGACCTTCCACGAGTGGACCTTCCGCCAGAGGGCACT 192
 Qy 70 LysAspPheLeu---IleTyrGluTyrArgAsnGlyHisGlnGluValGluSerPro--- 87
 Db 193 GAGGAGTTTGTCAAGATCTCTCGCGCTATGAGATGAGGTGTTCGACGCTGGAGGAGATGAG 252
 Qy 88 ---PheGlnGlyArgLeuGlnTrpAsnGlySer-----LysAspLeuGlnAspVal 103
 Db 253 CGCTTCGAGGCGCGCTGTGTGGATGGCAGCGGGGCGACCAAGACCTTGCAGATCTG 312
 Qy 104 SerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuTyrThrCysAsnValSer 123
 Db 313 TCTATCTTTCATCACCACCAATGTCACTACACCACTCGCGGCGACTACGAGTGCCAGTCTAC 372
 Qy 124 ArgGluPheGluPheGluAlaHisArgProPheValLysThrArgLeuIleProLeu 143
 Db 373 CGCTGTCTCTTCTTCGAAAACCTACGAGCACACACACCGCGTGTCAAGAAGATCCACATT 432
 Qy 144 ArgValThrGluGluAlaGlyGlu 151
 Db 433 GAGGTAGTGACAAA---GGTGAG 453

RESULT 12

US-10-276-774-718/c
 ; Sequence 718, Application US/10276774
 ; Publication No. US20040053245A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hyseq, Inc.
 ; APPLICANT: Tang, Y, Tom et al
 ; TITLE OF INVENTION: No. US20040053245A1el Nucleic Acids and Polypeptides
 ; FILE REFERENCE: 21273-030
 ; CURRENT APPLICATION NUMBER: US/10/276,774
 ; CURRENT FILING DATE: 2002-11-18
 ; PRIOR APPLICATION NUMBER: 09/560,875
 ; PRIOR FILING DATE: 2000-04-27
 ; PRIOR APPLICATION NUMBER: 09/496,914
 ; PRIOR FILING DATE: 2000-02-03
 ; NUMBER OF SEQ ID NOS: 2700
 ; SOFTWARE: Custom
 ; SEQ ID NO 718
 ; LENGTH: 407
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-276-774-718

Alignment Scores:
 Pred. No.: 6.16e-22 Length: 407
 Score: 234.00 Matches: 47
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 20.82% Indels: 0
 DB: 17 Gaps: 0

US-09-977-579-2 (1-215) x US-10-276-774-718 (1-407)

Qy 149 AlaGlyGluAspPheThrSerValSerGluIleMetMetTyrIleLeuValPhe 168
 Db 230 GCTGGAGAGGACTTCACCTCTGTGGTCTCAGAATCATGATGATACATCTCTGTCTTC 171
 Qy 169 LeuThrLeuTrpLeuLeuIleGluMetIleTyrCysTyrArgLysValSerLysAlaGlu 188
 Db 170 CTCACCTTGTGGTGTCTCATCGAGATGATATTTGCTACAGAAAGGTCTCAAAAGCGAA 111
 Qy 189 GluAlaAlaGlnGluAsnAla 195
 Db 110 GAGGCGACCCCAAGAAAACGCG 90
 RESULT 13
 US-10-205-331-111
 ; Sequence 111, Application US/10205331
 ; Publication No. US20040058326A1
 ; GENERAL INFORMATION:

```
/ APPLICANT: Warner-Lambert Company
/ APPLICANT: Lee, Kevin
/ APPLICANT: Dixon, Alistair
/ APPLICANT: Brooksbank, Robert
/ APPLICANT: Pincock, Robert
/ TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
/ FILE REFERENCE: WL-A-018199
/ CURRENT APPLICATION NUMBER: US/10/205,331
/ CURRENT FILING DATE: 2002-07-24
/ PRIOR APPLICATION NUMBER: GB 0118354.0
/ PRIOR FILING DATE: 2001-07-27
/ NUMBER OF SEQ ID NOS: 117
/ SOFTWARE: Patent In Ver. 2.1
/ SEQ ID NO 111
/ TYPE: DNA
/ LENGTH: 1029
/ ORGANISM: Rattus norvegicus
/ FEATURE:
/ OTHER INFORMATION: Schwann cell peripheral myelin
US-10-205-331-111

Alignment Scores:
Pred. No.: 4,13e-12 Length: 1029
Score: 167.00 Matches: 58
Percent Similarity: 43.81% Conservative: 34
Best Local Similarity: 27.62% Mismatches: 100
Query Match: 14.86% Indels: 18
DB: 17 Gaps: 7

US-09-977-579-2 (1-215) x US-10-205-331-111 (1-1029)
QY 11 AlaSerLeuValLeuLeuTyrTrpValSerValCysPheProValCysValGluValPro 30
DB 89 TCTTCTTGGTGTG-----TCCCAACCCCTGGCCATGTGGTTTACACGGAC 136
QY 31 SerGluThrGluAlaValGlnGlyAsnProMetLysLeuArgCysIleSerCysMetLys 50
DB 137 AGGGAAGTCTATGGTCTGGCTGCCCTCCAGGTACCCCTGCACCTGC---TCTTCTGGTCC 193
QY 51 ArgGluGluValGluAlaThrThrValValGluTyrPheTyrArgProGluGlyLys 70
DB 194 AGTGAATGGTCTCAGATGACATCTCTTTTACCTGGCGCTACACGCTGAAGAGGCCGA 253
QY 71 AspPheLeu---IleTyrGluTyrArgAsnGlyHisGlnGluValGlu-----SerPro 87
DB 254 GATGCCATTTCAATCTTCCACTATGCCAAGGGTCAACCTTACATCGATGAGTGGGGACC 313
QY 88 PheGlnGlyArgLeuGlnTrpAsnGlySerLysAspLeuGluAspValSerIleThrVal 107
DB 314 TTCAAGGAGCGCATCAGTGGGTAGGGAGCCCTAGCTGGAGATGGCTCCATTTGTCATA 373
QY 108 LeuAsnValThrLeuAsnAspSerGlyLeuTyrThrCysAsnValSerArgGluPheGlu 127
DB 374 CACAACCTAGACTACAGTCACAAGGCACTTTTACATGTGATGTCAAAAACCCACCGGAC 433
QY 128 PheGluAlaHisArgProPheValLysThrThrArgLeuLeuProLeuArgValThrGlu 147
DB 434 -----ATAGTGGGCAAGACGCTCTCAGGTCAAGCTCATGTCTATGTCTTTGAA 475
QY 148 GluAlaGlyGluAspPheThrSerValValSerGluLeuMetMetTyrIleLeuLeuVal 167
DB 476 AAAGTGCCACCTAGGTATGGGTGGTGGTGGAGCGCGTATCGTGGCATCTCTGGGGTG 535
QY 168 PheLeuThrLeuTrpLeuLeuLeuGluMetIle---TyrCysTyrArgLysValSerLys 186
DB 536 GTGCTTGTCTGTGTCTTCTTCTACCTGATCCGGTACTGTGGTGTGGCTGGCAGGAGGCT 595
QY 187 AlaGluGlu-----AlaAlaGlnGluAsnAlaSerAspTyrLeuAlaIleProSer 203
DB 596 GCCCTGCAGAGGAGGCTCAGTGCATGAGAGAGGGGAAATTTTCAAGATCTTCTTAAGGAC 655
QY 204 GluAsnLysGluAsnSerAlaValProVal 213
DB 111
QY 111
DB 111

Db 656 TCCTCGAAGCGCGGGCGGAGACGCGAGTG 685
RESULT 14
US-10-723-860-6471
/ Sequence 6471, Application US/10723860
/ Publication No. US20040253606A1
/ GENERAL INFORMATION:
/ APPLICANT: Aziz, Natasha
/ APPLICANT: Ginsburg, Wendy M.
/ APPLICANT: Zlotnik, Albert
/ TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
/ TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
/ FILE REFERENCE: 05882.0193.NPUS01
/ CURRENT APPLICATION NUMBER: US/10/723,860
/ CURRENT FILING DATE: 2003-11-26
/ PRIOR APPLICATION NUMBER: 60/429,739
/ PRIOR FILING DATE: 2002-11-26
/ NUMBER OF SEQ ID NOS: 8393
/ SOFTWARE: Patent In version 3.2
/ SEQ ID NO 6471
/ LENGTH: 3583
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (294)..(315)
/ OTHER INFORMATION: n is a, c, g, or t
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1696)..(1712)
/ OTHER INFORMATION: n is a, c, g, or t
US-10-723-860-6471

Alignment Scores:
Pred. No.: 4,58e-11 Length: 3583
Score: 165.50 Matches: 56
Percent Similarity: 47.11% Conservative: 26
Best Local Similarity: 32.18% Mismatches: 45
Query Match: 14.72% Indels: 48
DB: 18 Gaps: 7

US-09-977-579-2 (1-215) x US-10-723-860-6471 (1-3583)
QY 67 GluGlyGlyLysAspPheLeuLeuTyrGlu-TyrArgAsnGlyHisGlnGluValGluSe 86
DB 1851 GAAGGAGGTATTTCATTTTACAAATTGAAGAAACAGAGG-----CCTAGAGACATTGAGTC 1904
QY 86 rPropheGlnGlyArgLeu-----GlnTrp-----AsnGlySerLysAs 99
DB 1905 ACTTGGCAGAGTACACAGCTGGCGCAGTGGCAGAGCCAGCTTTCAACCTGAGCCTTCT 1964
QY 99 pheuGlnAspValSerIle-----ThrValLeuAsnValThrLeuAsnAspSerGlyLe 117
DB 1965 GGTTCGACAGCTACGGCTTTTAGGCACCGTGTGAGGGCTCCAG-AATGACACAGATGT 2023
QY 117 uTyrThrCysAsnValSerArgGluPheGluAlaHisArg-----132
DB 2024 GGCCTCGAGTTACACAGGAGCGGTGAGGGTGACCCCGCAGGCGCTTACCAAGGCT 2083
QY 133 -----ProPheValLysThrThrArgLeuLeuProLeu-----143
DB 2084 GGGTATTATATACAGTGCATACACAGGCGGAGGAGGTGAGCCACTCATCCAAGC 2143
QY 143 -----143
DB 2144 TCACACAGCAAGCTCACACAGCACTCAGGCTGTCTATGAGCCTGGGCTACCCCTTAAC 2203
QY 144 -ArgValThrGluGluAlaGlyGluAspPheThrSerValValSerGluLeuMetMetTy 163
DB 2204 CTGCTTGGCCCTGAGGAGCAACAGAGACATGCGATCCATCGTGTCTGAGATCATGTGA 2263
QY 163 rIleLeuLeuValPheLeuThrLeuTrpLeuLeuLeuGluMetIleTyrCysTyrArgLy 183
DB 111
QY 111
DB 111
```

Db 2264 TGTGCTCATTGTGCTGTGACCATATGCTGTGGCAGAGATGATTTACTGTACAGAA 2323
 Qy 183 sValSerLysAla---GluGluAlaAlaGlnGluAsnAla 195
 Db 2324 GATCGCTGCCGCCACGAGACTGCTGCACAGGAGATGCG 2363
 RESULT 15
 US-09-960-706-472
 ; Sequence 472, Application US/09960706
 ; Publication No. US20030134280A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Munger, William B.
 ; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplasia
 ; FILE OF INVENTION: Gene Expression Profiles
 ; FILE REFERENCE: 44921-5029-01US
 ; CURRENT APPLICATION NUMBER: US/09/960,706
 ; CURRENT FILING DATE: 2001-09-24
 ; PRIOR APPLICATION NUMBER: 60/223,323
 ; PRIOR FILING DATE: 2000-08-07
 ; PRIOR APPLICATION NUMBER: 09/873,319
 ; PRIOR FILING DATE: 2001-06-05
 ; NUMBER OF SEQ ID NOS: 1124
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 472
 ; LENGTH: 1929
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: Genbank Accession No. US20030134280A1 D10537
 US-09-960-706-472

Alignment Scores:
 Pred. No.: 3.31e-11 Length: 1929
 Score: 163.50 Matches: 54
 Percent Similarity: 45.55% Conservative: 33
 Best Local Similarity: 28.27% Mismatches: 81
 Query Match: 14.55% Indels: 23
 DB: 10 Gaps: 7
 US-09-977-579-2 (1-215) x US-09-960-706-472 (1-1929)

Qy 11 AlaSerLeuValLeuLeuTyrTrpValSerValCysPheProValCysValGluValPro 30
 Db 101 TCTTCTTTGGTGTG-----TCCCGCGCCCGAGCCCATCGTGGTTTACACCGAC 148
 Qy 31 SerGluThrGluAlaValGlnGlyAsnProMetLysLeuArgCysIleSerCysMetLys 50
 Db 149 AGGAGGTCCATGCTGTGGGCTCCCGGTGACCTGCACCTGC---TCTTCTGTCTC 205
 Qy 51 ArgGluGluValGluAlaThrThrValValGluTrpPheTyrArgProGluGlyGlyLys 70
 Db 206 AGTGAGTGGGTCTCAGATGACATCTCTTCCCTGGCGCTACCAGCCCGAGGAGGAGCAGA 265
 Qy 71 AspPheLeu---IleTyrGluTyrArgAsnGlyHisGlnGluValGlu-----SerPro 87
 Db 266 GATGCCATTTCGATCTTCCACTATGCCAGGACACCCCTACATTGACGAGGTGGGACC 325
 Qy 88 PheGlnGlyArgLeuGlnTrpAsnGlySerLysAspLeuGlnAspValSerIleThrVal 107
 Db 326 TTCAAAGAGCGCATCCAGTGGGTAGGGGACCTCGCTGGAGGATGGCTCCATTGTGCATA 385
 Qy 108 LeuAsnValThrLeuAsnAspSerGlyLeuTyrThrCysAsnValSerArgGluPheGlu 127
 Db 386 CACAACTAGACTACAGTACATGGCAGCTTCACTTGTGACGTCATAAAACCCCTCCAGAC 445
 Qy 128 PheGluAlaHisArgProPheValLysThrArgLeuLeuProLeuArgValThrGlu 147
 Db 446 -----ATAGTGGGCAGACACCTCTCAGGTACGCTGTATGCTTTTGA 487
 Qy 148 GluAlaGlyGluAspPheThrSerValSerGluIleMetMetTyrIleLeuVal 167
 Db 488 AAAGTCCCACTAGGTACGGGGTGTCTTCTGGAGCTGTGATCGGGGTGTCTCGGGGTG 547

Qy 168 PheLeuThrLeuTrpLeuLeuLeuGluMetIle---TyrCysTyr----- 181
 Db 548 GTGCTGTTGCTGTGCTGCTTTTCTACGTGGTTCGGTACTGCTGGCTACGACGAGCG 607
 Qy 182 -----ArgLysValSerLysAlaGluGlu 189
 Db 608 GCCCTGCAGAGAGAGGCTCAGTGTATGGAGAG 640
 Search completed: April 7, 2005, 17:21:00
 Job time : 508 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 7, 2005, 14:55:55 ; Search time 3093 Seconds
(without alignments)
2645.916 Million cell updates/sec

Title: US-09-977-579-2

Perfect score: 1124

Sequence: 1 MPANRFLPLASLVLIYVWS.....SDYLAIPSENKENSAPVVEE 215

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DSV=xlp
-Q/cgna2_1/USPTO_spool_p/US09977579/runat_06042005_163457_23604/app_query.fasta_1.391
-DB=EST -Qfmt=fastap -SUFFIX=ret -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-OUTFMT=pco -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=LOCAL
-USER=US09977579_CGNA_1_1_5180 @runat_06042005_163457_23604 -NCFU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST.*

1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | ID | Description |
|------------|-------|---------------|--------|-------------|----------------------|
| 1 | 1124 | 100.0 | 648 | 9 AY419145 | AY419145 Homo sapi |
| 2 | 1124 | 100.0 | 2555 | 3 CR609664 | CR609664 full-length |
| 3 | 1124 | 100.0 | 4052 | 3 HSM801563 | AL136589 Homo sapi |
| 4 | 1115 | 99.2 | 1062 | 5 BX420015 | BX420015 BX420015 |
| 5 | 1097 | 97.6 | 648 | 9 AY419147 | AY419147 Mus muscu |
| 6 | 1097 | 97.6 | 1359 | 3 AK076466 | AK076466 Mus muscu |
| 7 | 1097 | 97.6 | 3549 | 3 AK049747 | AK049747 Mus muscu |
| 8 | 1084 | 96.4 | 4149 | 3 AK049286 | AK049286 Mus muscu |
| 9 | 1063 | 94.6 | 636 | 5 BM933157 | BM933157 UI-M-BH3- |

| | | | | | |
|----|--------|------|------|---|-----------|
| 10 | 1061 | 94.4 | 975 | 5 | BX445002 |
| 11 | 1051 | 93.5 | 672 | 2 | BB614118 |
| 12 | 1042.5 | 92.7 | 4105 | 3 | EC058083 |
| 13 | 977 | 86.9 | 950 | 4 | BG294174 |
| 14 | 962 | 85.6 | 582 | 5 | BP200910 |
| 15 | 937 | 83.4 | 723 | 6 | CA749311 |
| 16 | 868.5 | 77.3 | 927 | 5 | BQ713131 |
| 17 | 818 | 72.8 | 584 | 5 | BP361278 |
| 18 | 809 | 72.0 | 986 | 5 | BUI18914 |
| 19 | 792 | 70.5 | 652 | 7 | CN219870 |
| 20 | 788.5 | 70.2 | 825 | 5 | BQ745919 |
| 21 | 774.5 | 68.9 | 846 | 6 | CD355879 |
| 22 | 760 | 67.6 | 700 | 4 | B1739617 |
| 23 | 747 | 66.5 | 582 | 5 | BP311484 |
| 24 | 747 | 66.5 | 825 | 5 | BUI368614 |
| 25 | 745 | 66.3 | 742 | 6 | CB526211 |
| 26 | 744 | 66.2 | 595 | 5 | BX368531 |
| 27 | 737 | 65.6 | 895 | 5 | BX743834 |
| 28 | 718 | 63.9 | 584 | 9 | AY419146 |
| 29 | 701 | 62.4 | 845 | 6 | CA327438 |
| 30 | 683 | 60.8 | 652 | 2 | BB652801 |
| 31 | 681 | 60.6 | 798 | 7 | CF290312 |
| 32 | 667.5 | 59.4 | 1069 | 5 | BM928131 |
| 33 | 645 | 57.4 | 450 | 6 | CB787935 |
| 34 | 622 | 55.3 | 734 | 7 | CK367344 |
| 35 | 618 | 55.0 | 823 | 7 | CO428886 |
| 36 | 606 | 53.9 | 712 | 6 | CB526257 |
| 37 | 600.5 | 53.4 | 901 | 7 | CF591773 |
| 38 | 583 | 51.9 | 780 | 5 | BQ770528 |
| 39 | 582 | 51.8 | 750 | 6 | CD349206 |
| 40 | 576.5 | 51.3 | 714 | 7 | CF531573 |
| 41 | 564 | 50.2 | 683 | 5 | BU492633 |
| 42 | 547.5 | 48.7 | 705 | 6 | CD349958 |
| 43 | 541.5 | 48.2 | 669 | 5 | BM951151 |
| 44 | 540 | 48.0 | 582 | 5 | BP202832 |
| 45 | 532 | 47.3 | 471 | 6 | CB732717 |

ALIGNMENTS

RESULT 1

AY419145

LOCUS

DEFINITION

AY419145 Homo sapiens HCM6793 gene, partial sequence, genomic survey sequence.

AY419145 GI:39775105

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 648)

Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferriterra, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.

Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios

Science 302 (5652), 1960-1963 (2003)

14671302

2 (bases 1 to 648)

Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferriterra, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.

Direct Submission

Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

These sequences were made by sequencing genomic exons and ordering them based on alignment.

Location/Qualifiers

1. .648

source

1. .648

1. .648

1. .648

1. .648

1. .648

1. .648

1. .648

1. .648

1. .648

1. .648

1. .648

1. .648

```
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<!--648
/locus_tag="HCM6793"

ORIGIN
Alignment Scores:
Pred. No.: 2,13e-126 Length: 648
Score: 1124.00 Matches: 215
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-977-579-2 (1-215) x AY419145 (1-648)
QY 1 MetProAlaPheAsnArgLeuPheProLeuAlaSerLeuValLeuIleTyrTrpValSer 20
Db 1 ATGCCTGCCTTCAATAGATTGTTTCCCTGGCTTCTCTGCTTATCTACTGGGTCAGT 60
QY 21 ValCysPheProValCysValGluValProSerGluThrGluAlaValGlnGlyAsnPro 40
Db 61 GTCGTCTCCCTGTGTGTGTGAAGTCCCTCGAGACGGAGCCGTGAGGGCAACCC 120
QY 41 MetLysLeuArgCysIleSerCysMetLysArgGluGluValGluAlaThrThrValVal 60
Db 121 ATGAAGCTCGCTGCATCTCTCGATGAAGAGAGAGAGAGGAGTGGAGGCCACACGGTGTG 180
QY 61 GluTrpPheTyrArgProGluGlyGlyAspPheLeuIleTyrGluTyrArgAsnGly 80
Db 181 GAATGTGTTCTCAGGCCCGAGGGCGGTAAAGATTTCCTTATTACAGATCGGAATGTC 240
QY 81 HisGlnGluValIleSerProPheGlnGlyArgLeuGlnTrpAsnGlySerLysAspLeu 100
Db 241 CACCAGAGGTGGAGAGCCCTTTCAGGGGCGCTGTCAGTGAATGGCAGAGACCTG 300
QY 101 GlnAspValSerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuTyrThrCys 120
Db 301 CAGGACGTGTCCATCCTGCTGCTCAACGTCACCTCTGAACGACTCTGGCCCTACACCTGC 360
QY 121 AsnValSerArgGluPheGluAlaHisArgProPheValLysThrThrArgLeu 140
Db 361 AATGTGTCGGGAGTTTGAGTTGAGGCGCATCGGCCCTTTGTGAAGACGACGGGCTG 420
QY 141 IleProLeuArgValThrGluGluAlaGlyGluAspPheThrSerValValSerGluIle 160
Db 421 ATCCCCCTAAGAGTCAACGAGGAGGCTGGAGAGGACTTCACCTCTGTGTCTCAGAAATC 480
QY 161 MetMetTyrIleLeuLeuValPheLeuThrLeuTrpLeuLeuIleGluMetIleTyrCys 180
Db 481 ATGATGTACATCCTTCTGTGCTCTCTCCTCCTGCTGCTCATCGAGATGATATATTC 540
QY 181 TyrArgLysValSerLysAlaGluGluAlaGlnGluAsnAlaSerAspTyrLeuAla 200
Db 541 TACAGAAAGTCTCAAAACCGAGAGGCGAGCCCAAGAAACGGCTGTGACTACCTTGCC 600
QY 201 IleProSerGluAsnLysGluAsnSerAlaValProValGluGlu 215
Db 601 ATCCCATCTGAGAACAAAGAGAACTCTGGGTACCACTGAGGAGAA 645

RESULT 2
CR609664
LOCUS
DEFINITION
full-length cDNA clone CS0DF023YA09 of Fetal brain of Homo sapiens
(human).
ACCESSION
CR609664
VERSION
CR609664.1 GI:50490471
KEYWORDS
HTC; CNSLT_CDNA.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 (bases 1 to 2555)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished
JOURNAL
REMARK
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue
Genoscope.
2 (bases 1 to 2555)
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE [E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
source
1..2555
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF023YA09"
/tissue_type="Fetal brain"
/plasmid="pCMVSPORT_6"

ORIGIN
Alignment Scores:
Pred. No.: 1,52e-125 Length: 2555
Score: 1124.00 Matches: 215
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

US-09-977-579-2 (1-215) x CR609664 (1-2555)
QY 1 MetProAlaPheAsnArgLeuPheProLeuAlaSerLeuValLeuIleTyrTrpValSer 20
Db 235 ATGCCTGCCTTCAATAGATTGTTTCCCTGGCTTCTCTGCTTATCTACTGGGTCAGT 294
QY 21 ValCysPheProValCysValGluValProSerGluThrGluAlaValGlnGlyAsnPro 40
Db 295 GTCGTCTCCCTGTGTGTGTGAAGTCCCTCGAGACGGAGCCGTGAGGGCAACCC 354
QY 41 MetLysLeuArgCysIleSerCysMetLysArgGluGluValGluAlaThrThrValVal 60
Db 355 ATGAAGCTCGCTGCATCTCTCGATGAAGAGAGAGAGTGGAGGCCACACGGTGTG 414
QY 61 GluTrpPheTyrArgProGluGlyGlyAspPheLeuIleTyrGluTyrArgAsnGly 80
Db 415 GAATGTGTTCTCAGGCCCGAGGGCGGTAAAGATTTCCTTATTACAGATCGGAATGTC 474
QY 81 HisGlnGluValIleSerProPheGlnGlyArgLeuGlnTrpAsnGlySerLysAspLeu 100
Db 475 CACCAGAGGTGGAGAGCCCTTTTCAGGGGCGCTGTCAGTGAATGGCAGAGGACCTG 534
QY 101 GlnAspValSerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuTyrThrCys 120
Db 535 CAGGACGTGTCCATCCTGCTGCTCAACGTCACCTCTGAACGACTCTGGCCCTACACCTGC 594
QY 121 AsnValSerArgGluPheGluPheGluAlaHisArgProPheValLysThrThrArgLeu 140
Db 595 AATGTGTCGGGAGTTTGAGTTGAGGCGCATCGGCCCTTTGTGAAGACGACGGGCTG 654
QY 141 IleProLeuArgValThrGluGluAlaGlyGluAspPheThrSerValValSerGluIle 160
Db 655 ATCCCCCTAAGAGTCAACGAGGAGGCTGGAGAGGACTTCACCTCTGTGTCTCAGAAATC 714
QY 161 MetMetTyrIleLeuLeuValPheLeuThrLeuTrpLeuLeuIleGluMetIleTyrCys 180
Db 715 ATGATGTACATCCTTCTGTGCTCTCTCCTCCTGCTGCTCATCGAGATGATATATTC 774
```



```

QY 181 TyrArgLysValSerLysAlaGluAlaGlnGluAlaSerAspTyrLeuAla 200
Db 775 TACAGAAAGGTCTCAAGAGCGAAGAGCGCCCAAGAAACGGCTCTGACTACCTTGCC 834
QY 201 IleProSerGluAsnLysGluAsnSerAlaValProValGluGlu 215
Db 835 ATCCCATCTGAGAACAGGAGAACTCTGCGGTACCACTGGAGGAA 879

RESULT 3
HSM801563
LOCUS HSM801563 4052 bp mRNA linear HTC 22-SEP-2004
DEFINITION Homo sapiens mRNA; cDNA DKFZp761F182 (from clone DKFZp761F182).
ACCESSION AL136589
VERSION AL136589.1 GI:13276680
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 4052)
AUTHORS Otenwaelder,B., Obermaier,B., Deutechenbaur,S., Schaipp,A.,
Mewes,H.W., Weil,B., Amid,C., Oeanger,A., Fobo,G., Han,M. and
Wiemann,S.
The German cDNA Consortium
Direct Submission
Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Medigenomix (Martinried/Germany) within the cDNA
sequencing consortium of the German Genome Project.
This clone (DKFZp761F182) is available at the RZPD Deutsches
Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp761F182
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.

FEATURES
source
1..4052
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFZp761F182"
/db_xref="taxon:9606"
/clone="DKFZp761F182"
/tissue_type="amygdala"
/clone_lib="761 (synonym: hamy2). Vector pSport1; host
DH10B; sites NotI + SalI"
/dev_stage="adult"
/notes="voltage-gated sodium channel beta-3 subunit"
1..4052
/gene="DKFZp761F182"
804..1451
/gene="DKFZp761F182"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAB66524.1"
/db_xref="GI:13276681"
/db_xref="GOA:Q9NY72"
/db_xref="UniProt/Swiss-Prot:Q9NY72"
/translations="MPAFNRLFPFLASLIYVWSVCFVPCVVEPSETEAVQGNPKMLR
CISMKRVEATVYVWFYRPEGKDFLIYVRNGHQBVEFPGRLQWNGSKLDQ
VSTLVNLTNLSGLTYCNVSRFEFEHAFHFPVKTTLPLRVTREAGDFTSVSEI
MYTLLVFLTLWLLIEMIYCYRKVSKAEAAQENASDYLAIPSENKENSAPVEE"

ORIGIN
Alignment Scores:
Pred. No.: 2,93e-125 Length: 4052
Score: 1124.00 Matches: 215
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 3 Gaps: 0

```

```

US-09-977-579-2 (1-215) x HSM801563 (1-4052)
QY 1 MetProAlaPheAsnArgLeuPheProLeuAlaSerLeuValLeuIleTyrTrpValSer 20
Db 804 ATGCTCGCTTCAATAGATTGTTTCCCTGGCTTCTCTCGTGCCTATCTACTGGTTCAGT 863
QY 21 ValCysPheProValCysValGluValProSerGluThrGluAlaValGlnGlyAsnPro 40
Db 864 GTCTGCTTCCCTGTGTGTGGAGTGCCTTCGGAGAGCGGAGCGCGTGCAGGGCAACCC 923
QY 41 MetLysLeuArgCysIleSerCysMetLysArgGluGluValGluAlaThrThrValVal 60
Db 924 ATGAAGCTGCGCTGCATCTCTGCATCAGAGAGAGAGAGGAGGAGGAGGAGGAGGAGG 983
QY 61 GluTrpPheTyrArgProGluGlyGlyAspPheLeuIleTyrGluTyrArgAsnGly 80
Db 984 GAATGGTTCTACAGGCGCGGCGGTAAGATTTCCTTATTTACGAGTATCGGAATGGC 1043
QY 81 HisGlnGluValGluSerProPheGlnGlyArgGluGlnTrpAsnGlySerLysAspLeu 100
Db 1044 CACCAGAGGTGGAGAGCCCTTTTCAGGGGCGCTTCAGTGGAGATGGCAGGAGACCTG 1103
QY 101 GlnAspValSerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuTyrThrCys 120
Db 1104 CAGGAGGTGTCATCACTGTCTCAACGTCACTCTGAACGACTCTGGCTCTACACCTGC 1163
QY 121 AsnValSerArgGluPheGluPheGluAlaHisArgProPheValLysThrArgLeu 140
Db 1164 AATGTGTCCCGGAGTTTGTAGTTTGTAGCGCATCGGCCCTTTGTGAAGACGACGCGGTG 1223
QY 141 IleProLeuArgValThrGluGluAlaGlyGluAspPheThrSerValValSerGluIle 160
Db 1224 ATCCCTTAAGATCCACCGAGAGGCTGGAGAGACTTCACCTCTGTGTCTCAGAAATC 1283
QY 161 MetMetTyrIleLeuLeuValPheLeuThrLeuTrpLeuLeuIleGluMetIleTyrCys 180
Db 1284 ATGATGTATACCTCTCTGTGTCTTCTCCTCACCTTGTGGCTGTCTATCGAGATATATTC 1343
QY 181 TyrArgLysValSerLysAlaGluAlaGlnGluAlaSerAspTyrLeuAla 200
Db 1344 TACAGAAAGGTCTCAAGAGCGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1403
QY 201 IleProSerGluAsnLysGluAsnSerAlaValProValGluGlu 215
Db 1404 ATCCCATCTGAGAACAGGAGAACTCTGCGGTACCACTGGAGGAGGAA 1448

RESULT 4
BX420015 1062 bp mRNA linear EST 01-MAY-2004
LOCUS BX420015 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
DEFINITION CS0DF023YA09 5-PRIME, mRNA sequence.
ACCESSION BX420015
VERSION BX420015.2 GI:46929710
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1062)
AUTHORS Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 13, 2003 this sequence version replaced gi:30646738.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.

```

This sequence belongs to sequence cluster 6147.r
 For more information about this cluster, see
<http://www.genoscope.cns.fr/cdna?s=CSODF023AA05QP1&c=6147.r>.

FEATURES
 source
 1..1062
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CSODF023YA09"
 /tissue type="FETAL BRAIN"
 /dev stage="fetal"
 /clone lib="Homo sapiens FETAL BRAIN"
 /notes="Organ: brain; Vector: pCMVSPORT.6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

Alignment Scores:
 Pred. No.: 5,42e-125 Length: 1062
 Score: 1115.00 Matches: 214
 Percent Similarity: 99.53% Conservative: 0
 Best Local Similarity: 99.53% Mismatches: 1
 Query Match: 99.20% Indels: 0
 DB: 5 Gaps: 0

US-09-977-579-2 (1-215) x BX420015 (1-1062)

QY 1 MetProAlaPheAsnArgLeuPheProLeuAlaSerLeuValLeuIleTyrTrpValSer 20
 |||||
 DB 234 ATGCCTGCCCTCAATAGATTGTTCCCTCGGCTTCTCTCGTGTATCTACTGGGTCAGT 293
 |||||

QY 21 ValCysPheProValCysValGluValProSerGluThrGluAlaValGlnGlyAsnPro 40
 |||||
 DB 294 GTCGTCTCCCTGTGTGTGGAGTGCCTCGGAGCGGAGCCGTGAGGGCAACCCC 353
 |||||

QY 41 MetLysLeuArgCysIleSerCysMetLysArgGluGluValGluAlaThrValVal 60
 |||||
 DB 354 ATGAAGCTGGCTGCATCTCTCGATGAAGAGAGAGAGGTGGAGGCCACCGTGGTG 413
 |||||

QY 61 GluTrpPheTyrArgProGluGlyGlyLysAspPheLeuIleTyrGluTyrArgAsnGly 80
 |||||
 DB 414 GAATGTTCTACAGGCCCGAGGCGGTAAAGATTCTCTTATTACGATTCGGAATGCG 473
 |||||

QY 81 HisGlnGluValGluSerProPheGlnGlyArgLeuGlnTrpAsnGlySerLysAspLeu 100
 |||||
 DB 474 CACCAGGAGTGGAGAGCCCTTTTCAGGGCGGCTGCGTGGAAATGGCAGGACCTG 533
 |||||

QY 101 GlnAspValSerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuTyrThrCys 120
 |||||
 DB 534 CAGGACGTGTCCATCTGTGTCTCAACGTCACCTCGAACGACTCTGGCCCTCTACACCTGC 593
 |||||

QY 121 AsnValSerArgGluPheGluAlaHisArgProPheValLysThrArgLeu 140
 |||||
 DB 594 ATGTGTCGGGAGTTGAGTTGAGGCGCATCGGCCCTTTGTGAAGACGACCGGCTG 653
 |||||

QY 141 IleProLeuArgValThrGluAlaGlyGluAspPheThrSerValValSerGluIle 160
 |||||
 DB 654 ATCCCTTAAGAGTCAACCGAGGAGGTGGAGGAGCTTCACCTCTGTGGTCTCAGAAATC 713
 |||||

QY 161 MetMetTyrIleLeuValPheLeuThrLeuTyrLeuIleGluMetIleTyrCys 180
 |||||
 DB 714 ATGATGTACATCTCTGTGTCTCTCCCTACCTGTGTGGCTGCTCATCGAGATGATATTGC 773
 |||||

QY 181 TyrArgLysValSerLysAlaGluAlaGlnGluAsnAlaSerAspTyrLeuAla 200
 |||||
 DB 774 TACAGAAAGGTCTCAAAAGCCGAGGAGGAGGCCCAAGAAACCGCTCTGACTACCTTGGC 833
 |||||

QY 201 IleProSerGluAsnLysGluAsnSerAlaValProValGluGlu 215
 |||||
 DB 834 ATCSCATCTGAGAACCAAGAGAACTCTGGCGTACCAGTGGAGGA 878
 |||||

RESULT 5

AY419147
 LOCUS
 DEFINITION
 Mus musculus HCM6793 gene, VIRTUAL TRANSCRIPT, partial sequence,
 genomic survey sequence.
 AY419147 648 bp DNA linear GSS 12-DEC-2003
 AY419147.1 GI:39775107
 GSS.
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 648)
 Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Perriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
 Adams,M.D. and Cargill,M.
 Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios
 Science 302 (5652), 1960-1963 (2003)
 PUBLISHED
 14671302
 REFERENCE 2 (bases 1 to 648)
 Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Perriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
 Adams,M.D. and Cargill,M.
 Direct Submission
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 These sequences were made by sequencing genomic exons and ordering
 them based on alignment.
 FEATURES
 Location/Qualifiers
 1..648
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 -1..->648
 /locus_tag="HCM6793"
 gene
 ORIGIN
 Alignment Scores:
 Pred. No.: 4,22e-123 Length: 648
 Score: 1097.00 Matches: 210
 Percent Similarity: 97.67% Conservative: 0
 Best Local Similarity: 97.67% Mismatches: 5
 Query Match: 97.60% Indels: 0
 DB: 9 Gaps: 0

US-09-977-579-2 (1-215) x AY419147 (1-648)

QY 1 MetProAlaPheAsnArgLeuPheProLeuAlaSerLeuValLeuIleTyrTrpValSer 20
 |||||
 DB 1 ATGCCTGCCCTCAACAGATTGCTTCCCTAGCTTCTCTAGTGTCTATCTACTGGGTCAGA 60
 |||||

QY 21 ValCysPheProValCysValGluValProSerGluThrGluAlaValGlnGlyAsnPro 40
 |||||
 DB 61 GTCGTCTCCCTGTGTGTAGAGTACCTCGGAGACAGACCCGTGAGGGCAATTCC 120
 |||||

QY 41 MetLysLeuArgCysIleSerCysMetLysArgGluGluValGluAlaThrValVal 60
 |||||
 DB 121 ATGAAGCTGAGATGCATCTCTCGATGAAGAGGAGGTGGAGGCCACCTGTAGTG 180
 |||||

QY 61 GluTrpPheTyrArgProGluGlyLysAspPheLeuIleTyrGluTyrArgAsnGly 80
 |||||
 DB 181 GAGTGGTCTTACAGGCCCTGAGGCGGTAAAGATTCTCTTATATATGAGTATCGAAATGGC 240
 |||||

QY 81 HisGlnGluValGluSerProPheGlnGlyArgLeuGlnTrpAsnGlySerLysAspLeu 100
 |||||
 DB 241 CACCAGAGGTGGAGAGCCCTTCCAGGTGCTGTCAGTGGAGTGGAGCAAGACCTG 300
 |||||

QY 101 GlnAspValSerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuTyrThrCys 120
 |||||
 DB 301 CAGGACGTATCCATCTGTTCTCAATGTCACTCTGAATGACTCTGGCCTCTACACATGT 360
 |||||

| | | | |
|------------|--|---|-----------------|
| QY | 121 | AsnValSerArgGluPheGluAlaHisArgProPheValIleThrThrArgLeu | 140 |
| Db | 361 | AAATGTCACAGGAGTTTGGTTTGAAGCACCAGGCTTTTGTGAAGACACAGACTA | 420 |
| QY | 141 | IleProLeuArgValThrGluGluAlaGlyGluAspPheThrSerValSerGluIle | 160 |
| Db | 421 | ATACCCCTGCGAGTCACTGAAGAGCGGAGAGACTTCACTCCGTCGGTCTCGGAATC | 480 |
| QY | 161 | MetMetTyrIleLeuLeuValPheLeuThrIleuTrpLeuLeuIleGluMetIleTyrCys | 180 |
| Db | 481 | ATGATGATACATCCCTCCCTGGCTTCTCCTCACTTGGCTGTTTATTTGAGATGATCTATTC | 540 |
| QY | 181 | TyrArgLysValSerLysAlaGluGluAlaGlnGluAsnAlaSerAspTyrLeuAla | 200 |
| Db | 541 | TACAGAAAGTCTCTAGGCGGAGAGGAGCTCAGGAAATGGCTGTGACTACCTTGCT | 600 |
| QY | 201 | IleProSerGluAsnLysGluAsnSerAlaValProValGluGlu | 215 |
| Db | 601 | ATCCCTTCAGAGCAAGGAGAACTCTGTGTGCTACCCCGTGAGGAA | 645 |
| RESULT 6 | | | |
| AK076466 | | | |
| LOCUS | | | |
| DEFINITION | AK076466 | 1359 bp mRNA linear | HTC 03-APR-2004 |
| | | Mus musculus 0 day neonate head cDNA, RIKEN full-length enriched | |
| | | library, clone:4833414B02 product:VOLTAGE-GATED SODIUM CHANNEL | |
| | | BETA-3 SUBUNIT (SODIUM CHANNEL BETA 3 SUBUNIT) homolog [Rattus | |
| | | norvegicus], full insert sequence. | |
| ACCESSION | AK076466 | | |
| VERSION | AK076466.1 | GI:26345409 | |
| KEYWORDS | HTC; CAP trapper. | | |
| SOURCE | Mus musculus (house mouse) | | |
| ORGANISM | Mus musculus | | |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | |
| | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | |
| REFERENCE | 1 | Carninci, P. and Hayashizaki, Y. | |
| AUTHORS | | | |
| TITLE | | High-efficiency full-length cDNA cloning | |
| JOURNAL | | Meth. Enzymol. 303, 19-44 (1999) | |
| MEDLINE | 99279253 | | |
| PUBMED | 10349636 | | |
| REFERENCE | 2 | | |
| AUTHORS | | | |
| TITLE | | Normalization and subtraction of cap-trapper-selected cDNAs to | |
| | | prepare full-length cDNA libraries for rapid discovery of new genes | |
| JOURNAL | | Genome Res. 10 (10), 1617-1630 (2000) | |
| MEDLINE | 20499374 | | |
| PUBMED | 11042159 | | |
| REFERENCE | 3 | | |
| AUTHORS | | | |
| TITLE | | Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., | |
| | | Konno, H., Akiyama, J., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M., | |
| | | Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., | |
| | | Yamanoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., | |
| | | Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., | |
| | | Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., | |
| | | Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. | |
| | | RIKEN integrated sequence analysis (RISA) system--384-format | |
| | | sequencing pipeline with 384 multicapillary sequencer | |
| JOURNAL | | Genome Res. 10 (11), 1757-1771 (2000) | |
| MEDLINE | 20530913 | | |
| PUBMED | 11076861 | | |
| REFERENCE | 4 | | |
| AUTHORS | | | |
| TITLE | | The RIKEN Genome Exploration Research Group Phase II Team and the | |
| | | FANTOM Consortium. | |
| | | Functional annotation of a full-length mouse cDNA collection | |
| JOURNAL | | Nature 409, 685-690 (2001) | |
| MEDLINE | 11076861 | | |
| PUBMED | 11076861 | | |
| REFERENCE | 5 | | |
| AUTHORS | | | |
| TITLE | | The FANTOM Consortium and the RIKEN Genome Exploration Research | |
| | | Group Phase I & II Team. | |
| | | Analysis of the mouse transcriptome based on functional annotation | |
| | | of 60,770 full-length cDNAs | |
| JOURNAL | | Nature 420, 563-573 (2002) | |

REFERENCE

AUTHORS

6 (bases 1 to 1359)

Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,
 Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
 Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
 Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
 Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,
 Koya, S., Kurihara, C., Matsuura, T., Miyazaki, A., Nishi, K.,
 Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
 Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
 Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,
 Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,
 Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I., Yasunishi, A.,
 Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.jp,
 URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
 Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.jp/

URL: http://fantom.gsc.riken.jp/

Location/Qualifiers

FEATURES

source

1. .1359
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM DB:4833414B02"
 /db_xref="taxon:10090"
 /clone="4833414B02"
 /tissue_type="head"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="0 day neonate"
 57. .704

CDS

/note="unnamed protein product; VOLTAGE-GATED SODIUM
 CHANNEL BETA-3 SUBUNIT (SODIUM CHANNEL BETA 3 SUBUNIT)
 homolog [Rattus norvegicus] (SPTR|Q9JK00, evidence: FASTY,
 99.5%ID, 100%length, match=646)
 putative"
 /codon_start=1
 /protein_id="BAC36356.1"
 /db_xref="GI:26345410"

/translation="MPAFNRLPLASLVLYVVRVCFVCPVEPSETEAVQGNMMLR
 CISCMEKEVEATTVVWFYRPGGKDFLIYVRNGHVEVSPQGRLOWNGSKDLQ
 VSTVLNVLNDSGLYTCNVSRPEPEHPPVKTTLIPLRVTTEAGEDPTSVVSEI
 MMILLVFLTLMDFIEMICYRKVSKAEAAQENASDYLAIPEKNSVVPVEE"

ORIGIN

Alignment Scores:

Pred. No.: 1-22e-122 Length: 1359
 Score: 1097.00 Matches: 210
 Percent Similarity: 97.67% Conservative: 0
 Best Local Similarity: 97.67% Mismatches: 5
 Query Match: 97.60% Indels: 0
 DB: 3 Gaps: 0

US-09-977-579-2 (1-215) x AK076466 (1-1359)

QY 1 MetProAlaPheAsnArgLeuPheProLeuAlaSerLeuValIleTyrTrpValSer 20

Db 57 ATGCCTGCCCTCAACAGATTGCTTCCCTAGCTCTCTAGTGTCTATCTACTGGTTCAGA 116

QY 21 ValCysPheProValCysValGluValProSerGluThrGluAlaValGlnGlyAsnPro 40

Db 117 GTCTGCTCTCCTGTGTGTAGAGTACCCCTCGAGAGACAGACCGGTGCGAGCAATTCC 176


```
Percent Similarity: 97.67%      Conservative: 0
Best Local Similarity: 97.67%    Mismatches: 5
Query Match: 97.60%             Indels: 0
DB: 3                           Gaps: 0

US-09-977-579-2 (1-215) x AK049747 (1-3549)

QY 1 MetProAlaPheAsnArgLeuPheProLeuAlaSerLeuValLeuIleTyrTrpValSer 20
DB 232 ATGCCTGCCTTCACACAGATTGCTCCCTAGCTTCTCTAGTGTCTCATCTACTGGTTCAGA 231
QY 21 ValCysPheProValCysValGluValProSerGluThrGluAlaValGlnGlyAsnPro 40
DB 292 GTCTGCTCCCTGCTGTGTAGAGTACCTCGGAGACAGAACGCGGCGGCAATTCC 351
QY 41 MetLysLeuArgCysIleSerCysMetLysArgGluValGluAlaThrValVal 60
DB 352 ATGAAGCTGAGATGCATCTCTCGATGAAGAGGAGGAGGTGGAGGCCACCATGTAGTG 411
QY 61 GluTrpPheTyrArgProGluGlyGlyLysAspPheLeuIleTyrGluTyrArgAsnGly 80
DB 412 GAGTGGTTCACAGGCTGAGGCGGTAAAGATTCTTATATATGATGATCGAATGGC 471
QY 81 HisGlnGluValGluSerProPheGlnGlyArgLeuGlnTrpAsnGlySerLysAspLeu 100
DB 472 CACCAGAGGTGGAGAGCCCTTCCAAGTGTCTGTCAGTGGAAATGGGAGCAAGACCTG 531
QY 101 GluAspValSerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuTyrThrCys 120
DB 532 CAGAGGATATCCATCACTGTTCTCAATGTCACTCTGAATCACTCTGGCCCTTACACATGT 591
QY 121 AsnValSerArgGluPheGluPheGluAlaHisArgProPheValLysThrThrArgLeu 140
DB 592 AATGTGTCAGAGGAGTTGAGTTCGAGACACACGGCCCTTTGTGAAGACACACAGACTA 651
QY 141 IleProLeuArgValThrGluGluAlaGlyGluAspPheThrSerValValSerGluIle 160
DB 652 ATACCCCTGCGAGTCACTGAAGAGCGGAGAGACTTCACCTCCGTGGTCTCGGAATC 711
QY 161 MetMetTyrIleLeuLeuValPheLeuThrLeuTrpLeuLeuIleGluMetIleTyrCys 180
DB 712 ATGATGTACATCTCTCTGGTCTTCTCCTCACCCTTGGCTGTTATTGATGATGATCTATTGC 771
QY 181 TyrArgLysValSerLysAlaGluGluAlaGlnGluAsnAlaSerAspTyrLeuAla 200
DB 772 TACAGAAAGTCTCTAAGCGCGAGAGGCGAGCTCAGGAAATGGCTGTGACTACCTTGCT 831
QY 201 IleProSerGluAsnLysGluAsnSerAlaValProValGluGlu 215
DB 832 ATCCCTTCAGAGAACAGGAGAACTCTGTGTACCCGTCGGAGGAA 876

RESULT 8
LOCUS AK049286
DEFINITION AK049286 4149 bp mRNA linear HTC 03-APR-2004
MUS MUSCULUS ES CELLS cDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY,
CLONE: C330019103 product: VOLTAGE-GATED SODIUM CHANNEL BETA-3
SUBUNIT (SODIUM CHANNEL BETA 3 SUBUNIT) homolog [Rattus
norvegicus], full insert sequence.
ACCESSION AK049286
VERSION AK049286.1 GI:26093400
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4
The RIKEN Genome Exploration Research Group Phase II Team and the
PANTOM Consortium
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The PANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 4149)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/.
FEATURES
source
1. .4149
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:C330019103"
/db_xref="taxon:10090"
/clone="C330019103"
/cell_type="ES cells"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
misc_feature
410..1058
/notes="VOLTAGE-GATED SODIUM CHANNEL BETA-3 SUBUNIT (SODIUM
CHANNEL BETA 3 SUBUNIT) homolog [Rattus norvegicus]
(SPTR|Q9JUK0, evidence: FASTY, 99.5%ID, 100%length),
```

```

match=646)
putative"
4129..4134
/note="putative"
4149
/note="putative"

ORIGIN
Alignment Scores:
Pred. No.:      2,32e-120      Length:      4149
Score:          1084.00      Matches:      210
Percent Similarity: 97.22%      Conservatives: 0
Best Local Similarity: 97.22%      Mismatches: 5
Query Match:      96.44%      Indels:      1
DB:               3          Gaps:      0

US-09-977-579-2 (1-215) x AK049286 (1-4149)
QY 1 MetProAlaPheAsnArgLeuPheProLeuAlaSerLeuValLeuIleTyrTrpValSer 20
DB 410 ATGCCTGCCCTTCAACAGATTGCTTCCCTAGCTTCTCTAGTCTCTATCTACTGGGTCTAGA 469
QY 21 ValCysPhe-ProValCysValGluValProSerGluThrGluAlaValGlnGlyAsnPr 40
DB 470 GTCTGCTTNCCTGTGTGTGTAGAGTACCTTCGGAGACAGAGCCGTGCAGGGCAATTC 529
QY 40 oMetLysLeuArgCysIleSerCysMetLysArgGluGluValGluAlaThrValVal 60
DB 530 CATGAAGCTGAGATGATCTCTCGCATGAAGAGGAGGAGGTGGAGGCCACCACTGTACT 589
QY 60 lGluTrpPheTyrArgProGluGlyGlyAspPheLeuIleTyrGluTyrArgAsnG1 80
DB 590 GGAGTGGTCTACAGCCCTCAGAGCGCGTAAAGATTTCCTATATATATAGTATCGAAATGG 649
QY 80 yHisGlnGluValGluSerProPheGlnGlyArgLeuGlnTrpAsnGlySerLysAspLe 100
DB 650 CCACAGAGGTGGAGAGCCCTTCAAGTGTCTGCAAGTGTGAGAGTGAATGGAGCAAGACT 709
QY 100 uGlnAspValSerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuTyrThrCy 120
DB 710 GCAGGAGGTATCCATCACTGTTCTCAATGTCACTCTGAATGACTCTGGGCTCTACACATG 769
QY 120 sAsnValSerArgGluPheGluPheGluAlaHisArgProPheValLysThrThrArgLe 140
DB 770 TAAATGTGTCCAGGAGTTTGAGTTCGAACACACACCGCCCTTTGTGAACACACACAGACT 829
QY 140 uIleProLeuArgValThrGluAlaGlyGluAspPheThrSerValValSerGluI1 160
DB 830 AATACCCCTGCGAGTCACTGAGAGGGGGAGAGACTTCACCTCGTGGTCTCGGAAT 889
QY 160 eMetMetTyrIleLeuLeuValPheLeuThrLeuThrLeuLeuIleGluMetIleTyrCy 180
DB 890 CATGATGATACATCTCTCTGCTCACTACCTTGTGGCTGTTATTAGATGATCTATTG 949
QY 180 sTyrArgLysValSerLysAlaGluAlaGlnAlaGlnGluAsnAlaSerAspTyrLeuAl 200
DB 950 CTACAGAAAGGTCTCTAAGCGCGAAGAGAGCTCAGGAAATGGCTTGACTACTCTTGC 1009
QY 200 alleleProSerGluAsnLysGluAsnSerAlaValProValGluGlu 215
DB 1010 TATCCCTTCAGAGACACAGAGACTCTGTGTGATACCCGTGGAGGAA 1055

RESULT 9
BM933157
LOCUS
DEFINITION
UI-M-BH3-bseq-d-04-0-UI.r1 NIH BMAP M.S4 Mus musculus cDNA clone
UI-M-BH3-bseq-d-04-0-UI 5', mRNA sequence.
BM933157
ACCESSION
BM933157.1 GI:19392309
VERSION
BM933157.1
KEYWORDS
Mus musculus (house mouse)
SOURCE
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 636)
 Bonaldo,M.P., Lennon,G. and Soares,M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 Genome Res. 6 (9), 791-806 (1996)
 97044477
 8889548
 Contact: Chin, H
 National Institute of Mental Health
 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
 20892-9643, USA
 Tel: 301 443 1706
 Fax: 301 443 9890
 Email: mEST@mail.nih.gov
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 Seq primer: M13 REVERSE.

Location/Qualifiers
 1. 636
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="CS7BL/6J"
 /db_xref="taxon:10090"
 /clone="UI-M-BH3-bseq-d-04-0-UI"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NIH_BMAP_M.S4"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not 1; Site 2: Eco RI; The
 NIH_BMAP_M.S4 library is a subtracted library of a series,
 ultimately derived from a mixture of individually tagged
 normalized libraries from ten regions of the mouse brain
 (cerebellum, brain stems, olfactory bulbs, hypothalamus,
 cortex, amygdala, basal ganglia, pineal gland, striatum,
 hippocampus) after a series of subtractions to reduce the
 representation of cDNAs from which ESTs had already been
 generated. The following serially subtracted libraries
 were generated in this process: NIH_BMAP_M.S4,
 NIH_BMAP_M.S3.3, NIH_BMAP_M.S3.2, NIH_BMAP_M.S3.1,
 NIH_BMAP_M.S2, NIH_BMAP_M.S1. The subtracted library
 (NIH_BMAP_M.S4) was constructed as follows: PCR amplified
 cDNA inserts from NIH_BMAP_M.S3.3, NIH_BMAP_M.S3.2, and
 NIH_BMAP_M.S3.1 clones from which 3' ESTs had been derived
 was used as a driver in a hybridization with a pool of
 the NIH_BMAP_M.S3.3, NIH_BMAP_M.S3.2, and NIH_BMAP_M.S3.1
 libraries in the form of single-stranded circles. The
 remaining single-stranded circles (subtracted library)
 was purified by hydroxyapatite column chromatography,
 converted to double-stranded circles and electroporated
 into DH10B bacteria (Life Technologies) to generate the
 NIH_BMAP_M.S4 library. This procedure has been previously
 described (Bonaldo, Lennon and Soares, Genome Research
 6:791-806, 1996)"

ORIGIN

Alignment Scores:
 Pred. No.: 5,82e-119 Length: 636
 Score: 1063.00 Matches: 203
 Percent Similarity: 97.13% Conservatives: 0
 Best Local Similarity: 97.13% Mismatches: 6
 Query Match: 94.57% Indels: 0
 DB: 5 Gaps: 0

US-09-977-579-2 (1-215) x BM933157 (1-636)

QY 1 MetProAlaPheAsnArgLeuPheProLeuAlaSerLeuValLeuIleTyrTrpValSer 20
 DB 7 ATGCCTGCCCTTCAACAGATTGCTTCCCTAGCTTCTCTAGTCTCTATCTACTGGGTCTAGA 66

```

Qy 21 ValCysPheProValCysValGluValProSerGluThrGluAlaValGlnGlyAsnPro 40
Db 67 GTCTGCTTCCCTGTGTGTAGAGTACCCCTCGGAGACAGAGCGGTGCGAGGCAATCC 126
Qy 41 MetLysLeuArgCysIleSerCysMetLysArgGluGluValGluAlaThrThrValVal 60
Db 127 ATGAAGCTGATGATCATCTCTCTGATGAAGAGGAGGAGGTGGAGGCCACCACTGTAGTG 186
Qy 61 GluTrpPheTyrArgProGluGluGlyLysAspPheLeuIleTyrGluTyrArgAsnGly 80
Db 187 GAGTGGTTCACAGCCTCAGGCGGTAAAGATTTCCTTATATAGATATCGAAATGGC 246
Qy 81 HisGlnGluValGluSerProPheGlnGlyArgLeuGlnTrpAsnGlySerLysAspLeu 100
Db 247 CACCAGGAGGTGGAGAGCCCTTCCAAAGTCTGTCTGAGTGAATGGAGCAAGACCTG 306
Qy 101 GlnAspValSerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuTyrThrCys 120
Db 307 CAGGACGTATCCATCACTCTCTCAATGTCACTCTGAATGACTCTGGCCTCTACACATGT 366
Qy 121 AsnValSerArgGluPheGluPheGluAlaHisArgProPheValLysThrArgLeu 140
Db 367 AATGTGTCAGGAGGTGTAGTTCGAAGCACACCGGCCCTTTTGAAGACCACAGACTA 426
Qy 141 IleProLeuArgValThrGluGluAlaGlyLysPheThrSerValSerGluIle 160
Db 427 ATACCCCTGCGAGTCACTGAAGAGCGGAGAGACTTCACTCCGTGGTCTCGGAATC 486
Qy 161 MetMetTyrIleLeuLeuValPheLeuThrLeuTrpLeuIleGluMetIleTyrCys 180
Db 487 ATGATGTACATCTCTCTGCTCTTCTCACTCTGCTGTGGTGTATTTAGATGATCTATTC 546
Qy 181 TyrArgLysValSerLysAlaGluGluAlaGlnGluAsnAlaSerAspTyrLeuAla 200
Db 547 TACAGAAAGTCTCTAGTTCGAGAGCGAGCTCAGGAAATGCGTCTGACTACCTTGCT 606
Qy 201 IleProSerGluAsnLysGluAsnSer 209
Db 607 ATCCCTTCAGAGAACACAGAACTCT 633

```

```

RESULT 10
BX445002
LOCUS
DEFINITION
BX445002 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
CS0DF023YA09 5-PRIME, mRNA sequence.
BX445002
BX445002.2 GI:47009181
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 975)
Li,W.B., Gruber,C., Jesses,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 15, 2003 this sequence version replaced gi:30782286.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 6147.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS1AF006ZE05QP1&c=6147.r.
FEATURES
source
Location/Qualifiers
1..975
/organism="Homo sapiens"
/mol_type="mRNA"

```

```

/db xref="taxon:9606"
/clone="CS0DF023YA09"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/notes="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."
ORIGIN
Alignment Scores:
Pred. No.: 1.88e-118 Length: 975
Score: 1061.00 Matches: 209
Percent Similarity: 97.67% Conservativity: 1
Best Local Similarity: 97.21% Mismatches: 5
Query Match: 94.40% Indels: 1
DB: 5 Gaps: 0
US-09-977-579-2 (1-215) x BX445002 (1-975)
Qy 1 MetProAlaPheAsnArgLeuPheProLeuAlaSerLeuValLeuIleTyrTrpValSer 20
Db 235 ATGCCTGCTTCAATAGATTGTTTCCCTGGCTTCTCTCGTGTATCTACTGGGTCACT 294
Qy 21 ValCysPheProValCysValGluValProSerGluThrGluAlaValGlnGlyAsnPro 40
Db 295 GTCTGCTTCCCTGTGTGTGGAGTCCCTCGGAGCGAGGCCGTGCGAGGCACACCCC 354
Qy 41 MetLysLeuArgCysIleSerCysMetLysArgGluGluValGluAlaThrThrValVal 60
Db 355 ATGAAGCTGCGTCACTCTCTGTCATGAGAGAGAGAGGTGGAGGCCACACCGTGGTG 414
Qy 61 GluTrpPheTyrArgProGluGlyGlyLysAspPheLeuIleTyrGluTyrArgAsnGly 80
Db 415 GAATGGTTCACAGGCCCGGCGGTAAAGATTCTCTTATTACAGAGTATCGGAATGGC 474
Qy 81 HisGlnGluValGluSerProPheGlnGlyArgLeuGlnTrpAsnGlySerLysAspLeu 100
Db 475 CACCAGGAGTGGAGAGCCCTTTTCAGGGCGCTGAGTGAATGGAGCAGCAGGACCTG 534
Qy 101 GlnAspValSerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuTyrThrCys 120
Db 535 CAGGACGTGCTCATCACTGTGCTCAACGTCACTGTAACGACTCTGGCCTCTACACCARC 594
Qy 121 AsnValSerArgGluPheGluAlaHisArgProPheValLysThrArgLeu 140
Db 595 AAWAAGTCCCGGAGTTTGTAGTTTGTAGCGCATCGGCCCTTTGTGAAGACGACGCGCTG 654
Qy 141 IleProLeuArgValThrGluGluAlaGlyLysPheThrSerValValSerGluIle 160
Db 655 ATCCCTTAAAGTACACCGAGGAGGTGGAGAGACTTCACTCTGTGGTCTCAGAAATC 714
Qy 161 MetMetTyrIleLeuLeuValPheLeuThrLeuTrpLeuIleGluMetIleTyrCys 180
Db 715 ATGATGTATACATCTCTGCTCTCTCACTCTGCTGCTCATCGAGATGATATATTC 774
Qy 181 TyrArgLysValSerLysAlaGluGluAlaGlnGluAsnAlaSerAspTyrLeuAla 200
Db 775 TACAGAAAGTCTCAAAAGCCGAGAGGAGGCCCAAGAAACCGCTCTGACTACCTTGCC 834
Qy 201 IleProSerGluAsnLysGluAsnSerAlaValProValGluGlu 215
Db 835 ATCCATCTGAGAACAGGAGAC-TCTGCGGTATCACTGGAGAT 878
RESULT 11
BB614118
LOCUS
DEFINITION
BB614118 RIKEN full-length enriched, 0 day neonate head Mus
musculus cDNA clone 4833414B02 5', mRNA sequence.
ACCESSION
BB614118
VERSION
BB614118.1 GI:16454578

```


| | |
|---|---|
| REFERENCE AUTHORS | 1 (bases 1 to 4105) Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Sanchez, A., Helton, M., Madan, A., Young, A.C., Rodriguez, S., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalios, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002) 22388257 |
| JOURNAL MEDLINE PUBMED | 12477932 |
| REFERENCE | 2 (bases 1 to 4105) |
| AUTHORS | Strausberg, R. |
| JOURNAL | Submitted (08-SEP-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA |
| REMARK COMMENT | NIH-MGC Project URL: http://mgc.nci.nih.gov Contact: MGC help desk Email: cgabs-remail.nih.gov Tissue Procurement: Dr. Jim Lin, University of Iowa cDNA Library Preparation: M. Bento Soares, University of Iowa cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada info@bcsc.bc.ca Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Lilaea Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smalios, Michael Smith, Lorraine Spence, Jeff Scott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra. |
| FEATURES source | Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAK Plate: 126 Row: 1 Column: 19 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 23943798 This clone has the following problem: frame shifted. Location/Qualifiers 1..4105 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6" /db_xref="taxon:10090" /clone="IMAGE:6826414" /tissue_type="Brain, mouse, 13.5,14.5,16.5,17.5 dpc" /clone_lib="NIH_BMAP_FY0" /lab_host="DH10B" /note="Vector: pYX-ASC" |
| ORIGIN | Alignment Scores: Pred. No.: 2,67e-115 Length: 4105 Score: 1042.50 Matches: 208 Percent Similarity: 96.74% Conservative: 0 Best Local Similarity: 96.74% Mismatches: 6 |
| Query Match: | 92.75% |
| DB: | 3 Indels: 3 Gaps: 1 |
| US-09-977-579-2 (1-215) x BC058083 (1-4105) | |
| QY | 1 MetProAlaPheAenArgLeuPheProLeuAlaSerLeuValLeuIleTyrTrpValSer 20 |
| Db | 356 ATGCTGCTTCAACAGATTGCTTCCCTAGTCTCTCTAGTGTCTCATCTACTGGGA--- 411 |
| QY | 21 ValCysPheProValCysValGluValProSerGluThrGluAlaValGlnGlyAsnPro 40 |
| Db | 412 GTCTGCTTCCCTGTGTGTAGTAAGTACCTTCGAGACAGAACCGGTGCGAGGCAATTC 471 |
| QY | 41 MetLysLeuArgCysIleSerCysMetLysArgGluValGluAlaThrThrValVal 60 |
| Db | 472 ATGAGCTGAGATGATCTCTCTGATGAGAGGAGGAGGTGGAGGCCACCATGTAGTG 531 |
| QY | 61 GluTrpPheTyrArgProGluGlyGlyLysAspPheLeuIleTyrGluTyrArgAsnGly 80 |
| Db | 532 GAGTGGTCTTACAGGCTGAGGCGGTAAAGATTTCCTTATATATAGTAGTATCGAAATGGC 591 |
| QY | 81 HisGlnGluValGluSerProPheGlnGlyArgLeuGlnTrpAsnGlySerLysAspLeu 100 |
| Db | 592 CACCAGAGGTGGAGAGCCCTTCCAAAGTCTGTCAGTGGATGGAGCAA-GACCTG 650 |
| QY | 101 GlnAspValSerIleThrValLeuAenValThrLeuAenAspSerGlyLeuTyrThrCys 120 |
| Db | 651 CAGAGATATCAATCACTGTTCTCAATGTCATCTGANTGACTCTGGCCTCTACATGT 710 |
| QY | 121 AsnValSerArgGluPheGluPheGluAlaHisArgProPheValLysThrThrArgLeu 140 |
| Db | 711 AATGTGTCCAGGGAGTTTGAAGTTCGAAGCACACCGGCCCTTTGTGAAGACCACCAAGACTA 770 |
| QY | 141 IleProLeuArgValThrGluGluAlaGlyGluAspPheThrSerValValSerGluIle 160 |
| Db | 771 ATACCCCTCGAGTCACTGAAGAGCGGGAGAGACTTCACCTCCGCGGTCTCGGAATC 830 |
| QY | 161 MetMetTyrIleLeuLeuValPheLeuThrLeuTyrLeuLeuIleGluMetIleTyrCys 180 |
| Db | 831 ATGATGTACATCTCTCTGCTCTCCACCTTGTGGCTGTTTATGAGATGATCTATTGC 890 |
| QY | 181 TyrArgLysValSerLysAlaGluAlaGlnGluAenAlaSerAspTyrLeuAla 200 |
| Db | 891 TACAGAAAAGGTCTCTAAGCGGAGAGGAGCTCAGGAAAATGCGTCTGACTACCTTGTCT 950 |
| QY | 201 IleProSerGluAsnLysGluAenSerAlaValProValGluGlu 215 |
| Db | 951 ATCCCTTCAGAGAACCAAGGAGAACTCTGTGTGTTACCCGTGGAGGAA 995 |
| RESULT 13 | |
| LOCUS | BG294174 950 bp mRNA linear EST 21-FEB-2001 |
| DEFINITION | 602391245F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:4503250 5', mRNA sequence. |
| ACCESSION | BG294174 |
| VERSION | BG294174.1 GI:13054543 |
| KEYWORDS | EST. |
| SOURCE | Mus musculus (house mouse) |
| ORGANISM | Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. |
| REFERENCE | 1 (bases 1 to 950) NIH-MGC http://mgc.nci.nih.gov/ National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgabs-remail.nih.gov Tissue Procurement: The Cepko Laboratory cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: |

http://image.llnl.gov
Plate: LLMW10373 row: i column: 11
High quality sequence stop: 643.
Location/Qualifiers
1..950
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:4503250"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_94"
/note="Organ: eye; Vector: pCMV-SPORT6; Site 1: Noti;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 3.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

FEATURES

source

ORIGIN

Alignment Scores:
Pred. No.: 3,27e-108 Length: 950
Score: 977.00 Matches: 198
Percent Similarity: 93.95% Conservative: 4
Best Local Similarity: 92.09% Mismatches: 11
Query Match: 86.92% Indels: 4
DB: 4 Gaps: 2

US-09-977-579-2 (1-215) x BG294174 (1-950)

QY 1 MetProAlaPheAsnArgLeuPheProLeuAlaSerLeuValLeuIleTyrTrpValSer 20
DB 87 ATGCCTGCTTCAACAGATTGCTTCCCTAGCTTCTTAGTGTCTATCTACTGGGTGAGA 146
QY 21 ValCysPheProValCysValGluValProSerGluThrGluAlaValGlnGlyAsnPro 40
DB 147 GTCCTGCTCCCTGTGTGTAGAGTACCTCGGAGACAGAGCCGTGAGGGCAATTC 206
QY 41 MetLysLeuArgCysIleSerCysMetLysArgGluValGluAlaThrThrValVal 60
DB 207 ATGAAGCTGAGATGATCTCTCCGATGAAGAGGAGGAGGTGGAGGCCACCACTGTAGT 266
QY 61 GluTyrPheTyrArgProGluGlyLysAspPheLeuIleTyrGluTyrArgAsnGly 80
DB 267 GAGTGGTCTACAGGCTGAGGCGGTAAAGATTCTTATATAGATTCGAAATGGC 326
QY 81 HisGlnGluValGluSerProPheGlnGlyArgLeuGlnTrpAsnGlySerLysAspLeu 100
DB 327 CACCAGAGGTGGAGAGCCCTTCCAAAGTCTGTCAGTGGAAATGGGAGCAAGACCTG 386
QY 101 GlnAspValSerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuTyrThrCys 120
DB 387 CAGGACGTATCCATCTCTTCAATGTCATCTCTGAATGACTCTGGCTCTACACATGT 446
QY 121 AsnValSerArgGluPheGluPheGluAlaHisArgProPheValLysThrThrArgLeu 140
DB 447 AATGTGTCAG-GAGTTTGAGTTGAGACACACCGGCCCTTTGTGAGACCAACAGACTA 505
QY 141 IleProLeuArgValThrGluGluAlaGlyGluAspPheThrSerValValSerGluIle 160
DB 506 ATACCCCTCGAGTCACTCAAGAGGCGGAGAGAACTTCACCTCCGTGCTCGGAATC 565
QY 161 MetMetTyrIleLeuLeuValPheLeuThrLeuTyrLeuIleGluMetIleTyrCys 180
DB 566 ATGATGTATACCTCTCTGCTCTCTCCCTACCTTGTGGCTCTTTATTGAGATGATCTATTC 625
QY 181 TyrArgLysValSerLysAlaGluGluAlaAlaGlnGluAsnAlaSerAspTyrLeuAla 200
DB 626 TACAGAAAGGTTCTTAAGCCCGAGAGGAGCT-CAGAGAAATGCGTCTGACTACTTG--- 581
QY 201 IleProSerGluAsnLysGluAsnSerAlaValProValGluGlu 215
DB 682 TATCCCTTAGAGAACAGGCAA---CCTGTGTGATCCGTGCGATAAG 723

RESULT 14

BP200910

LOCUS

DEFINITION

BP200910 Sugano cDNA library, amygdala Homo sapiens cDNA clone

AMR03894, mRNA sequence.

ACCESSION

BP200910

VERSION

BP200910.1 GI:52050356

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 582)

AUTHORS

Suzuki, I., Yamashita, R., Shiota, M., Sakakibara, Y., Chiba, J.,

Mizushima-Sugano, J., Nakai, K. and Sugano, S.

TITLE

Sequence comparison of human and mouse genes reveals a homologous

block structure in the promoter regions

JOURNAL

Genome Res. 14 (9), 1711-1718 (2004)

COMMENT

Contact: Yutaka Suzuki

Department of Virology

Institute of Medical Science, University of Tokyo

4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan

Email: ysuzuki@ims.u-tokyo.ac.jp.

FEATURES

source

1..582

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="AMR03894"

/tissue_type="amygdala"

/clone_lib="Sugano cDNA library, amygdala"

ORIGIN

Alignment Scores:

Pred. No.: 1.1e-106 Length: 582

Score: 962.00 Matches: 182

Percent Similarity: 99.45% Conservative: 0

Best Local Similarity: 99.45% Mismatches: 1

Query Match: 85.59% Indels: 0

DB: 5 Gaps: 0

US-09-977-579-2 (1-215) x BP200910 (1-582)

QY 2 ProAlaPheAsnArgLeuPheProLeuAlaSerLeuValLeuIleTyrTrpValSerVal 21
DB 34 CTGCTCTTCAATAGATTTCTCCCTGGCTTCTCTGCTCTTCTACTGGGTGAGTGC 93
QY 22 CysPheProValCysValGluValProSerGluThrGluAlaValGlnGlyAsnProMet 41
DB 94 TGCTTCCCTGTGTGTGGAGTGCCTCGGAGACGGAGCGCTGCAGGCAACCCCATG 153
QY 42 LysLeuArgCysIleSerCysMetLysArgGluValGluAlaThrThrValValGlu 61
DB 154 AAGCTCGCTGCATCTCTGCATGAAGAGAGAGGAGGTGGAGGCCACCGGTGGTGA 213
QY 62 TrpPheTyrArgProGluGlyGlyLysAspPheLeuIleTyrGluTyrArgAsnGlyHis 81
DB 214 TGGTCTACAGCCCGAGGCGGTAAAGATTTCCTTATTACAGTATCGAATGGCCAC 273
QY 82 GlnGluValGluSerProPheGlnGlyArgLeuGlnTrpAsnGlySerLysAspLeuGln 101
DB 274 CAGGAGGTGGAGAGCCCTTTCAGGGGCGCTTCAGTGAATGGCAGCAAGACCTGCAG 333
QY 102 AspValSerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuTyrThrCysAsn 121
DB 334 GACGTGTCCATCACTGTCTCAACGTCACCTCGAACGACTCTGGCTCTACACCTCAAT 393
QY 122 ValSerArgGluPheGluAlaHisArgProPheValLysThrThrArgLeuIle 141
DB 394 GTGTCGCGGAGTTGAGTTGAGGCGCATCGCCCTTTGTGAAGACGACGGGCTGATC 453
QY 142 ProLeuArgValThrGluGluAlaGlyGluAspPheThrSerValValSerGluIleMet 161

Db 454 CCCTAAGAGTCAACGAGGAGGCTGAGAGGAGCTTCACTCTCTGTGGTCTCAGAAATCATG 513
 QY 162 MetTyrIleLeuLeuValPheLeuThrLeuTrpLeuIleLeuGluMetIleTyrCysTyr 181
 Db 514 ATGTATCATCTTCTGGTCTCCTCACCCTTGCTGTGGTCTCATCAGATGATATATTGCTAC 573
 QY 182 ArgLysVal 184
 Db 574 AGAAGGTC 582

RESULT 15

CA749311 723 bp mRNA linear EST 09-JUL-2003
 LOCUS UI-M-FY0-cdd-k-17-0-UI.r1 NIH BMAP_FY0 Mus musculus cDNA clone
 DEFINITION IMAGE:6831594 5', mRNA sequence.

ACCESSION CA749311
 VERSION
 KEYWORDS
 SOURCE CA749311.1 GI:25570984
 EST.

ORGANISM Mus musculus (house mouse)

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 723)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Dr. Jim Lin, University of Iowa

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

<http://image.lnlni.gov>

This clone was contributed by the Brain Molecular Anatomy Project

(BMAP)

Seq primer: pyX-5.

Location/Qualifiers

1. 723
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAG:6831594"
 /tissue_type="whole brain"
 /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
 /lab_host="DH10B (TI phage resistant)"
 /clone_lib="NIH BMAP FY0"

/notes="Organ: Brain; Vector: pyX- Asc; Site 1: Ecor I;
 Site 2: Not I; The library was constructed according
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was
 size selected according to mRNA size fraction, ligated
 with Ecor I adaptor, digested with NotI and then cloned
 directionally into pyX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is AGGAGACAG. This library was created for the University
 Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
 Developing Mouse Nervous System', supported by National
 Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
 program coordinator."

ORIGIN

Alignment Scores:

Pred. No.: 1.7e-103 Length: 723
 Score: 937.00 Matches: 180
 Percent Similarity: 97.30% Conservative: 0
 Best Local Similarity: 97.30% Mismatches: 4
 Query Match: 83.36% Indels: 1
 DB: 6 Gaps: 0

US-09-977-579-2 (1-215) x CA749311 (1-723)

QY 1 MetProAlaPheAsnArgLeuPheProLeuAlaSerLeuValLeuIleTyrTrpValSer 20
 Db 170 ATGCTGCGCTTCAACAGATTGCTTCCCTTAGCTTCTTAGTGTCTCATCTACTGCGGTGAGA 229
 QY 21 ValCysPheProValCysValGluValProSerGluThrGluAlaValGlnGlyAsnPro 40
 Db 230 GTCTGCTTCCCTGTGTGTAGAGTACCTTCGAGACAGAGCCGCTGAGGGCAATTC 289
 QY 41 MetLysLeuArgCysIleSerCysMetLysArgGluGluValGluAlaThrValVal 60
 Db 290 ATCAAGCTGAGATGCATCTCTCATGAAGAGGAGGAGGTGGAGGCCACCACTGTAGTG 349
 QY 61 GluTrpPheTyrArgProGluGlyGlyLysAspPheLeuIleTyrGluTyrArgAsnGly 80
 Db 350 GAGTGGTTCTACAGGCTGAGGCGGTGAAGATTCTCTATATATGATGATCGAAATGGC 409
 QY 81 HisGlnGluValGluSerProPheGlnGlyArgLeuGlnTrpAsnGlySerLysAspLeu 100
 Db 410 CACCAGAGGTGGAGAGCCCTTCCAAGTCTGTCGAGTGGATGGGAGCAAGACCTG 469
 QY 101 GlnAspValSerIleThrValLeuAsnValThrLeuAsnAspSerGlyLeuTyrThrCys 120
 Db 470 CAGGACGTATCCATCACTGTTCTCAATGTCAATGTCAATGACTCTGGCCTCTACACATGT 529
 QY 121 AsnValSerArgGluPheGluAlaHisArgProPheValLysThrThrArgLeu 140
 Db 530 AATGTGTCCAGGGAGTTGAGTTTCGAAGCACACCGGCCCTTTGTGAAGACCAAGACTA 589
 QY 141 IleProLeuArgValThrGluGlu-AlaGlyGluAspPheThrSerValValSerGluI 160
 Db 590 ATACCTCTGGAGTCACTGAGAGGCGGGAGAGACTTCACCTCCGTCTCGGAAT 649
 QY 160 eMetMetTyrIleLeuLeuValPheLeuThrLeuTrpLeuLeuGluMetIleTyrC 180
 Db 650 CATGATGTACATCCTCTGCTGCTCTCTCACTTGTGGCTGTTTATTGAGATGATCTATTG 709
 QY 180 sTyrArgLysVal 184
 Db 710 CTACAGAAAGGTC 722

Search completed: April 7, 2005, 17:09:38

Job time : 3104 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2005, 07:44:40 ; Search time 1135.27 Seconds
(without alignments)
10158.276 Million cell updates/sec

Title: US-09-977-579-4_COPY_1024_1261

Perfect score: 238

Sequence: 1 aacaggagcagtgtgacatg.....tgccagaactgagaagcgg 238

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_on.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------------|-------------|
| 1 | 238 | 100.0 | 1261 | 6 AR359850 | Sequence |
| 2 | 238 | 100.0 | 1261 | 6 AX039100 | Sequence |
| 3 | 238 | 100.0 | 1261 | 9 HSA243396 | Homo sapi |
| 4 | 236 | 99.2 | 5306 | 9 AB032984 | Homo sapi |
| 5 | 216 | 90.8 | 127347 | 2 AP000682 | Homo sapi |
| 6 | 216 | 90.8 | 144833 | 2 AC063921 | Homo sapi |
| 7 | 216 | 90.8 | 149800 | 2 AC021981 | Homo sapi |
| 8 | 216 | 90.8 | 172546 | 2 AC024604 | Homo sapi |
| 9 | 216 | 90.8 | 178169 | 9 AP002765 | Homo sapi |
| 10 | 216 | 90.8 | 178431 | 2 AP002749 | Homo sapi |
| 11 | 216 | 90.8 | 181471 | 9 AC069539 | Homo sapi |
| 12 | 203.6 | 85.5 | 3296 | 9 AB097521 | Macaca fa |
| 13 | 118 | 49.6 | 174285 | 2 AC145484 | Lemur cat |
| 14 | 108 | 45.4 | 66980 | 2 AC013796 | Homo sapi |
| 15 | 47.4 | 19.9 | 189486 | 2 AC114854 | Rattus no |
| 16 | 44.2 | 18.6 | 175316 | 5 BX649282 | Zebrafish |
| 17 | 44 | 18.5 | 182917 | 9 AL139383 | Human DNA |
| 18 | 43.8 | 18.4 | 178225 | 5 BX005286 | Zebrafish |
| 19 | 43.2 | 18.2 | 74549 | 9 HSH990015 | Human DNA |

| | | | | | | | |
|---|----|------|------|--------|----|----------|-----------|
| c | 20 | 43.2 | 18.2 | 139084 | 5 | AL935332 | Zebrafish |
| | 21 | 43.2 | 18.2 | 196687 | 10 | AC122466 | Mus muscu |
| | 22 | 43.2 | 18.2 | 199229 | 10 | AC120004 | Mus muscu |
| | 23 | 43.2 | 18.2 | 345383 | 2 | AC149590 | Mus muscu |
| c | 24 | 42.8 | 18.0 | 184751 | 5 | BX571948 | Zebrafish |
| c | 25 | 42.6 | 17.9 | 206768 | 2 | AC125951 | Rattus no |
| c | 26 | 42.6 | 17.9 | 228381 | 2 | AC103262 | Rattus no |
| c | 27 | 42.6 | 17.9 | 240645 | 2 | AC126633 | Rattus no |
| c | 28 | 42.4 | 17.8 | 61434 | 9 | AC007366 | Homo sapi |
| c | 29 | 42.4 | 17.8 | 93599 | 2 | AC084728 | Homo sapi |
| c | 30 | 42.4 | 17.8 | 118674 | 2 | AC023063 | Homo sapi |
| c | 31 | 42.4 | 17.8 | 174388 | 2 | BX842570 | Homo sapi |
| c | 32 | 42.4 | 17.8 | 195922 | 10 | AC140213 | Mus muscu |
| c | 33 | 42.4 | 17.8 | 204131 | 2 | AC132010 | Homo sapi |
| c | 34 | 42.2 | 17.7 | 129502 | 9 | HSJ726C3 | Human DNA |
| c | 35 | 41.8 | 17.6 | 159084 | 9 | AC138512 | Homo sapi |
| c | 36 | 41.8 | 17.6 | 205610 | 9 | AC134312 | Homo sapi |
| c | 37 | 41.6 | 17.5 | 196341 | 2 | BX293552 | Danio rer |
| c | 38 | 41.6 | 17.5 | 248002 | 5 | BX323588 | Zebrafish |
| c | 39 | 41.4 | 17.4 | 99961 | 9 | AL606510 | Human DNA |
| c | 40 | 41.4 | 17.4 | 226950 | 10 | AL671011 | Mouse DNA |
| c | 41 | 41.2 | 17.3 | 151804 | 2 | AC088161 | Mus muscu |
| c | 42 | 41.2 | 17.3 | 152684 | 2 | BX897684 | Danio rer |
| c | 43 | 41.2 | 17.3 | 157907 | 2 | CR847996 | Danio rer |
| c | 44 | 41.2 | 17.3 | 180745 | 9 | AL359183 | Human DNA |
| c | 45 | 41.2 | 17.3 | 205379 | 10 | AL591174 | Mouse DNA |

ALIGNMENTS

RESULT 1
AR359850
LOCUS AR359850 1261 bp DNA linear PAT 17-AUG-2003
DEFINITION Sequence 4 from patent US 6593565.
ACCESSION AR359850
VERSION AR359850.1 GI:33766660
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1261)
AUTHORS Heslin, P. and Lynam, N.R.
TITLE Vehicle interior rearview mirror assembly including an accessory-containing housing
JOURNAL Patent: US 6593565-A 4 15-JUL-2003;
FEATURES Location/Qualifiers
source 1. 1261
/organism="unknown"
/mol_type="genomic DNA"

| | | | | |
|-----------------------|-----------------|---|-----------|--------------|
| Query Match | 100.0%; | Score 238; | DB 6; | Length 1261; |
| Best Local Similarity | 100.0%; | Pred. No. 3.1e-59; | | |
| Matches 238; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |
| Qy | 1 | AACAGGAGCAGTGTGACATGAGGTGGCTGGAACACCTGAGGGACTGGACATCCCATGTC | 60 | |
| Db | 1024 | AACAGGAGCAGTGTGACATGAGGTGGCTGGAACACCTGAGGGACTGGACATCCCATGTC | 1083 | |
| Qy | 61 | AGCAATGTCAATGGCATCAGGAGGGCGGCCCAAGGGCCCATCGCTTCCCTTCATGCATC | 120 | |
| Db | 1084 | AGCAATGTCAATGGCATCAGGAGGGCGGCCCAAGGGCCCATCGCTTCCCTTCATGCATC | 1143 | |
| Qy | 121 | CATTGTTCTGTTCAATTCATTCATCCATATACCATCCCTGCTTGGCTTTCACCTCTGAC | 180 | |
| Db | 1144 | CATTGTTCTGTTCAATTCATTCATCCATATACCATCCCTGCTTGGCTTTCACCTCTGAC | 1203 | |
| Qy | 181 | TCCTTAACCTCCATCAGACCTCTACGACCACTAAGACTCTGCCAGAACTGAGAGCCGG | 238 | |
| Db | 1204 | TCCTTAACCTCCATCAGACCTCTACGACCACTAAGACTCTGCCAGAACTGAGAGCCGG | 1261 | |

20039618
10574461
PUBMED
REFERENCE
2 (bases 1 to 5306)
Ohara, O., Nagase, T. and Kikuno, R.
AUTHORS
TITLE
DIRECT SUBMISSION
JOURNAL
Submitted (04-OCT-1999) Osamu Ohara, Kazusa DNA Research Institute,
Laboratory of DNA Technology; 1532-3 Yana, Kisarazu, Chiba
292-0812, Japan (E-mail:cdnaiffo@kazusa.or.jp,
URL:http://www.kazusa.or.jp/huge/, Tel:+81-438-52-3913,
Fax:+81-438-52-3914)
COMMENT
Location/Qualifiers
1..5306
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="bj00081"
/sex="male"
/issue_type="brain"
/clone_lib="pBluescriptII SK plus"
/dev_stage="adult"
1..5306
/gene="KIAA1158"
1..695
/note="KIAA1158"
/note="Start codon is not identified."
/codon_start=3
/product="KIAA1158 protein"
/protein_id="BA086472.1"
/db_xref="GI:6330136"
/translation="RSWGSESLGSRQPKMPAFNRLPLASLVLIYVWVCFPVCEV
PSETAVQGNPKLRICSMKREVEATTVVWFYRGGKDFLIYRNGHGVESV
FQRLQNGSDLODSITLVNLTNLSGLYTCNVSRFEAFHRPFVTKTRILPVR
TEAGEDFTSVVSEIMVYLLVFLTWLLIEMIVCYRKVSKAEBAEENASDYLAIPS
ENKENSAPVVEE"
ORIGIN
Query Match 99.2%; Score 236; DB 9; Length 5306;
Best Local Similarity 100.0%; Pred. No. 1.1e-58;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACAGAGCAGGTGACATGAGTGGCTGACACACCTGAGGACTGGACATCCCATGTTTC 60
DB 696 AACAGAGCAGGTGACATGAGTGGCTGACACACCTGAGGACTGGACATCCCATGTTTC 755
QY 61 AGCAATGTCAATGGCATCAGGAGGGGCCCAAGGCCCTCGCTTCCCTTCATGCATC 120
DB 756 AGCAATGTCAATGGCATCAGGAGGGGCCCAAGGCCCTCGCTTCCCTTCATGCATC 815
QY 121 CATTTGTTCTGTTTCATTCATTCATCATCATCATCATCCCTGCTGAGCTTTTCACCTCTGAC 180
DB 816 CATTTGTTCTGTTTCATTCATTCATCATCATCATCCCTGCTGAGCTTTTCACCTCTGAC 875
QY 181 TCCTTAATCTCATCAGACCTCTACGCACCATTAAGACTCTGCGAGACTGAGAGCC 236
DB 876 TCCTTAATCTCATCAGACCTCTACGCACCATTAAGACTCTGCGAGACTGAGAGCC 931
RESULT 5
AP000682/c
LOCUS
DEFINITION
Homo sapiens chromosome 11 clone CMB9-32A1 map 11q24, WORKING DRAFT
SEQUENCE, 8 unordered pieces.
ACCESSION
AP000682.3 GI:9844967
VERSION
HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 127347)
Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
AUTHORS
TITLE
DIRECT SUBMISSION
JOURNAL
Submitted (08-NOV-1999) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/
Tel:+81-42-778-9923, Fax:+81-42-778-9924)
COMMENT
On Aug 18, 2000 this sequence version replaced gi:8118870.
----- Genome Center
Center: RIKEN Genomic Sciences Center(GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
----- Project Information
Center project name: HumDraft11
Center clone name: CMB9-32A1
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 124981 bases at least Q40
Consensus quality: 125815 bases at least Q30
Consensus quality: 126286 bases at least Q20
Insert size: 126647; sum-of-contigs
Quality coverage: 12.71x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of
8 contigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs N, but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved.
1 33532 contig of 33532 bp in length
33633 59223 contig of 25591 bp in length
59324 83007 contig of 23684 bp in length
83108 100196 contig of 17089 bp in length
100297 112313 contig of 12017 bp in length
112414 121043 contig of 8630 bp in length
121144 125917 contig of 4774 bp in length
126018 127347 contig of 1330 bp in length.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 33532 contig of 33532 bp in length
33633 59223 contig of 25591 bp in length
59324 83007 contig of 23684 bp in length
83108 100196 contig of 17089 bp in length
100297 112313 contig of 12017 bp in length
112414 121043 contig of 8630 bp in length
121144 125917 contig of 4774 bp in length
126018 127347 contig of 1330 bp in length.
Location/Qualifiers
1..127347
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
FEATURES
source

```

/map="11q24"
/clone="CWB9-32A1"
1. .33532
/note="assembly_fragment"
33633..59223
/note="assembly_fragment"
59324..83007
/note="assembly_fragment"
83108..100196
/note="assembly_fragment clone_end:SP6 vector_side:left"
100297..112313
/note="assembly_fragment"
112414..121043
/note="assembly_fragment clone_end:T7 vector_side:right"
121144..125917
/note="assembly_fragment"
126018..127347
/note="assembly_fragment"

ORIGIN
Query Match          90.8%; Score 216; DB 2; Length 127347;
Best Local Similarity 100.0%; Pred. No. 6.9e-53;
Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 21 AGGTGCGCTGAACACCTGAGGAGCTGGACATCCCATGTTTCAGCAATGTCATGGCATCAG 80
|||||
DB 8751 AGGTGCGCTGAACACCTGAGGAGCTGGACATCCCATGTTTCAGCAATGTCATGGCATCAG 8692
|||||
QY 81 GAGGGGCGCCCAAGGCGCCCATCGCTTCCTTCATGCATCCATGTTTCGTTTCATTCATT 140
|||||
DB 8691 GAGGGGCGCCCAAGGCGCCCATCGCTTCCTTCATGCATCCATGTTTCGTTTCATTCATT 8632
|||||
QY 141 CATCCATACATCAACCTGCTCTGAGCTTTCACCTCTGACTCCCTAACTCCATCAGACCT 200
|||||
DB 8631 CATCCATACATCAACCTGCTCTGAGCTTTCACCTCTGACTCCCTAACTCCATCAGACCT 8572
|||||
QY 201 CTACGGCACCATAAGACTCTGCCAGACTGAGAGCC 236
|||||
DB 8571 CTACGGCACCATAAGACTCTGCCAGACTGAGAGCC 8536

RESULT 6
AC063921 144833 bp DNA linear HTG 08-JAN-2003
LOCUS Homo sapiens chromosome 3 clone RP11-142P10, WORKING DRAFT
DEFINITION SEQUENCES, 13 unordered pieces.
ACCESSION AC063921
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 144833)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,
Alsbrooks,S.L., Anarstange,H.C., Are,J.R., Ayale,M., Banks,T.,
Barbaris,J., Benton,J., Bimge,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,B., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carton,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
Homsí,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,

```

```

Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louisedge,H.,
Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapa,P., Martin,R., Martindale,A., Martinez,E.,
Massey,B., Mawhinney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokwkw,S., Ogih,M., Okwuonu,G.,
Oragunye,N., Oviado,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Petars,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savary,G.,
Scherer,S., Scott,G., Shen,H., Shoostari,N., Sisson,I.,
Sodergren,B., Sonaik,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczkyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 144833)
Worley,K.C.
Direct Submission
Submitted (22-APR-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 144833)
Worley,K.C.
Direct Submission
Submitted (08-JAN-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Apr 28, 2002 this sequence version replaced gi:16117924.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: HAXY
Center clone name: RP11-142P10
----- Summary Statistics
Sequencing vector: M13;
Chemistry: Dye-terminator Big Dye: 86% of reads
Assembly program: Phrap: version 0.990329
Consensus quality: 127260 bases at least Q40
Consensus quality: 132735 bases at least Q30
Consensus quality: 136183 bases at least Q20
Estimated insert size: 139078; sum-of-contigs estimation
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 13 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 2437: contig of 2437 bp in length
* 2438 2537: gap of unknown length
* 2538 4812: contig of 2275 bp in length
* 4813 4912: gap of unknown length
* 4913 8488: contig of 3576 bp in length
* 8489 8589: gap of unknown length
* 8589 11163: contig of 2575 bp in length
* 11164 11263: gap of unknown length

```


Center code: GTC
Web site: <http://www.genomecorp.com/>
Contact: gtc-seqcenter@genomecorp.com
----- Project Information
Center project name: hg202
----- Summary Statistics

Sequencing vector: N/A
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 990315
Consensus quality: 152033 bases at least Q40
Consensus quality: 163535 bases at least Q30
Consensus quality: 165256 bases at least Q20
Insert size: 170446; sum-of-contigs
Quality coverage: 4.3x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 22 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1
* 1339: Contig of 1339 bp in length
* 1340
* 1439: Gap of unknown length
* 1440
* 2889: Contig of 1350 bp in length
* 2790
* 2889: Gap of unknown length
* 2890
* 4562: Contig of 1673 bp in length
* 4563
* 4662: Gap of unknown length
* 6344: Contig of 1682 bp in length
* 6345
* 6444: Gap of unknown length
* 9221: Contig of 2777 bp in length
* 9222
* 9321: Gap of unknown length
* 11997: Contig of 2676 bp in length
* 12097: Gap of unknown length
* 12098
* 14867: Contig of 2770 bp in length
* 14868
* 14967: Gap of unknown length
* 19919: Contig of 4952 bp in length
* 19920
* 20019: Gap of unknown length
* 20020
* 24125: Contig of 4106 bp in length
* 24126
* 24225: Gap of unknown length
* 24226
* 29444: Contig of 5219 bp in length
* 29445
* 29544: Gap of unknown length
* 29545
* 34568: Contig of 5024 bp in length
* 34569
* 34668: Gap of unknown length
* 39414: Contig of 4746 bp in length
* 39415
* 39514: Gap of unknown length
* 39515
* 45885: Contig of 6371 bp in length
* 45886
* 45985: Gap of unknown length
* 50997: Contig of 5012 bp in length
* 50998
* 51097: Gap of unknown length
* 51098
* 59143: Contig of 8046 bp in length
* 59144
* 59243: Gap of unknown length
* 59244
* 67779: Contig of 8536 bp in length
* 67780
* 67879: Gap of unknown length
* 67880
* 76429: Contig of 8550 bp in length
* 76430
* 76529: Gap of unknown length
* 76530
* 86704: Contig of 10175 bp in length
* 86705
* 86804: Gap of unknown length
* 86805
* 103147: Contig of 16343 bp in length
* 103148
* 103247: Gap of unknown length
* 103248
* 116343: Contig of 13096 bp in length
* 116344
* 116443: Gap of unknown length
* 116444
* 135318: Contig of 18875 bp in length
* 135319
* 135419: Gap of unknown length
* 135419
* 172546: Contig of 37128 bp in length.

FEATURES
source

1. .172546
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="Chromosome 10"
/clone="RP11-331G19"

ORIGIN
/clone_lib="RPC1-11"

Query Match 90.8%; Score 216; DB 2; Length 172546;
Best Local Similarity 100.0%; Pred. No. 6.8e-53;
Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 AGGTGGCCTGAACACCTGAGGGAGCTGGACATCCCATGTTCCAGCAATGTCATGCGCATCAG 80
Db 170778 AGGTGGCCTGAACACCTGAGGGAGCTGGACATCCCATGTTCCAGCAATGTCATGCGCATCAG 170719
Qy 81 GAGGGGGCCCCCAAGGGCCCCCATCGCTTCCCTTCATGCATCCATTGTTCTGTTTCATTCATT 140
Db 170718 GAGGGGGCCCCCAAGGGCCCCCATCGCTTCCCTTCATGCATCCATTGTTCTGTTTCATTCATT 170659

Qy 141 CATCCATACATCCACCTGCTCTGAGCTTTCACCTCTGACTCCCTAACTCATCAGACCT 200
Db 170658 CATCCATACATCCACCTGCTCTGAGCTTTCACCTCTGACTCCCTAACTCATCAGACCT 170599
Qy 201 CTACGGCACCATAAGACTCTGCCAGAACTGAGAAGCC 236
Db 170598 CTACGGCACCATAAGACTCTGCCAGAACTGAGAAGCC 170563

RESULT 9

AP002765/c

LOCUS

DEFINITION

complete sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1. .178169

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="11"

/map="11q"

/clone="RP11-634B22"

ORIGIN

Query Match

Best Local Similarity

Matches

216; Conservative

0; Mismatches

0; Indels

0; Gaps

0;

Qy

21

AGGTGGCCTGAACACCTGAGGGAGCTGGACATCCCATGTTCCAGCAATGTCATGCGCATCAG 80

Db

87137

AGGTGGCCTGAACACCTGAGGGAGCTGGACATCCCATGTTCCAGCAATGTCATGCGCATCAG 87078

Qy

81

GAGGGGGCCCCAAGGGCCCCCATCGCTTCCCTTCATGCATCCATTGTTCTGTTTCATTCATT 140

Db

87077

GAGGGGGCCCCAAGGGCCCCCATCGCTTCCCTTCATGCATCCATTGTTCTGTTTCATTCATT 87018

| | | | |
|--|-------------|-------------------------------------|-------|
| Db | 59095 | CTACGACCATAAGACTCTGCCGAACCTGAGAAGCC | 59060 |
| RESULT 12 | | | |
| AB097521 | | | |
| LOCUS | | | |
| DEFINITION | | | |
| AB097521 | | | |
| MACACA fascicularis brain cDNA clone; QmoA-13657, full insert sequence. | | | |
| AB097521.1 | GI:26449236 | | |
| OLIGO capping; ffs (full insert sequence). | | | |
| MACACA fascicularis (crab-eating macaque) | | | |
| MACACA fascicularis | | | |
| Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; | | | |
| Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; | | | |
| Cercopithecinae; Macaca. | | | |
| 1 | | | |
| Oada, N., Hida, M., Kusuda, J., Tanuma, R., Iseki, K., Hirata, M., Suto, Y., Hirai, M., Terao, K., Suzuki, Y., Sugano, S. and Hashimoto, K. | | | |
| Assignment of 118 novel cDNAs of cynomolgus monkey brain to human chromosomes | | | |
| Gene 275 (1), 31-37 (2001) | | | |
| 21458551 | | | |
| 11574149 | | | |
| 2 (bases 1 to 3296) | | | |
| Hashimoto, K., Osada, N., Hida, M., Kusuda, J. and Sugano, S. | | | |
| Direct Submission | | | |
| Submitted (05-DEC-2002) Katsuyuki Hashimoto, National Institute of Infectious Diseases, Division of Genetic Resources; 23-1, Toyama 1-chome, Shinjuku-ku, Tokyo 162-8640, Japan | | | |
| (E-mail: khashi@nih.go.jp, URL: http://www.nih.go.jp/yoken/genebank/, Tel: 81-3-5285-1111 (ex. 2120), Fax: 81-3-5285-1181) | | | |
| Lab host: TOPIO | | | |
| Vector: pME18S-FL3 (Acc.No. AB009864) | | | |
| R. Site1: DraIII (CACTGTGTG) | | | |
| R. Site2: DraIII (CACCATGGT) | | | |
| Description: 1st strand cDNA was primed with an oligo(dT) primer [ATGTGCCTTTTCTTTTTTTT]; double-stranded cDNA was synthesized using specific 5' and 3' primers and amplified by PCR. The PCR product was digested with SfiI and size selection was performed to exclude fragments <1.5kb. The SfiI-digested PCR product was cloned into distinct DraIII sites of pME18S-FL3. XhoI sites just outside the DraIII sites can be used to isolate the cDNA insert. Libraries were constructed by oligo-capping method. Custom primers used for sequencing (5' end primer [CTCTGCTCTAAAGCTCGG]; 3' end primer [CGACCTGCAGCTGAGCAC]). | | | |
| Location/Qualifiers | | | |
| 1..3296 | | | |
| /organism="Macaca fascicularis" | | | |
| /mol_type="mRNA" | | | |
| /db_xref="taxon:9541" | | | |
| /clone="QmoA-13657" | | | |
| /sex="male" | | | |
| /tissue.type="medulla oblongata" | | | |
| /clone.lib="macaque brain cDNA library QmoA" | | | |
| /dev_stage="adult" | | | |
| 6..853 | | | |
| /codon_start=1 | | | |
| /product="hypothetical protein" | | | |
| /protein_id="BAC41746.1" | | | |
| /db_xref="GI:26449237" | | | |
| translation="MPAFNRLLPPLSLVLIYASVCPCVCEVPSETEAVQGNPKMLR CISCWKREVEATTIVWFYFPEGKDFLIETYNHGEVESFPQRLQWNGSKDLD SVILNVLNDSGLYTCNVSRFEFAHRPFVKTLRIPLRVTEEAGEDFTSVSEI MMVIYLFLTLMTLEMIYCVRKSKAEAAQENASDYLAIPSENKENSAPVVEE" | | | |
| Query Match 85.5%; Score 203.6; DB 9; Length 3296; | | | |
| Best Local Similarity 94.5%; Pred. No. 4e-49; | | | |
| Matches 223; Conservative 0; Mismatches 9; Indels 4; Gaps 1; | | | |
| ORIGIN | | | |
| 1 AACAGGAGCAGTGTCACATGAGTGGCTGCAACCTGAGGACCTGGACATCCCATGTC | | | |
| 2 | | | |
| 3 | | | |
| 4 | | | |
| 5 | | | |
| 6 | | | |
| 7 | | | |
| 8 | | | |
| 9 | | | |
| 10 | | | |
| 11 | | | |
| 12 | | | |
| 13 | | | |
| 14 | | | |
| 15 | | | |
| 16 | | | |
| 17 | | | |
| 18 | | | |
| 19 | | | |
| 20 | | | |
| 21 | | | |
| 22 | | | |
| 23 | | | |
| 24 | | | |
| 25 | | | |
| 26 | | | |
| 27 | | | |
| 28 | | | |
| 29 | | | |
| 30 | | | |
| 31 | | | |
| 32 | | | |
| 33 | | | |
| 34 | | | |
| 35 | | | |
| 36 | | | |
| 37 | | | |
| 38 | | | |
| 39 | | | |
| 40 | | | |
| 41 | | | |
| 42 | | | |
| 43 | | | |
| 44 | | | |
| 45 | | | |
| 46 | | | |
| 47 | | | |
| 48 | | | |
| 49 | | | |
| 50 | | | |
| 51 | | | |
| 52 | | | |
| 53 | | | |
| 54 | | | |
| 55 | | | |
| 56 | | | |
| 57 | | | |
| 58 | | | |
| 59 | | | |
| 60 | | | |
| 61 | | | |
| 62 | | | |
| 63 | | | |
| 64 | | | |
| 65 | | | |
| 66 | | | |
| 67 | | | |
| 68 | | | |
| 69 | | | |
| 70 | | | |
| 71 | | | |
| 72 | | | |
| 73 | | | |
| 74 | | | |
| 75 | | | |
| 76 | | | |
| 77 | | | |
| 78 | | | |
| 79 | | | |
| 80 | | | |
| 81 | | | |
| 82 | | | |
| 83 | | | |
| 84 | | | |
| 85 | | | |
| 86 | | | |
| 87 | | | |
| 88 | | | |
| 89 | | | |

* This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 * 1 87364: contig of 87364 bp in length
 * 87365 87464: gap of unknown length
 * 87465 151123: contig of 63659 bp in length
 * 151124 151223: gap of unknown length
 * 151224 174285: contig of 23062 bp in length.

FEATURES

Location/Qualifiers
 1..174285
 /organism="Lemur catta"
 /mol_type="genomic DNA"
 /db_xref="taxon:9447"
 /clone="LB2-244B5"

ORIGIN

Query Match 49.6%; Score 118; DB 2; Length 174285;
 Best Local Similarity 81.0%; Pred. No. 5.2e-24;
 Matches 175; Conservative 0; Mismatches 35; Indels 6; Gaps 3;
 QY 21 AGGTGGCTGAACACCTGAGGACTGGACATCCCATCTTCAGCAATCTCAATGTCATCAG 80
 Db 3358 AGGTGGCTGAACACCTGAGGACTGGACATCCCATCTTCAGCAATCTCAATGTCATCAG 3300
 QY 81 GAGGCGCCCCCAAGGCGCCCATCGCTTCCTTCATCATCATCTTCATGTTCTGTTTCATTCATT 140
 Db 3299 GAGGCGCCCCCTCGGG-CCCATCTTCCTTCAGTGCCCATCTTCT-ATTTGTT 3245
 QY 141 CATCATACATCATCAGCTGCTTCAGCTTCACCTTCGACTCCCTAACTCATCAGACCT 200
 Db 3244 CATCATACATCTACCATCTCTGAGGCTCACCCTGACTGCTGCTAGTTCATCAGACCT 3185
 QY 201 CTAGCACCATAAGACTCTGCAGAACTGAGAGCC 236
 Db 3184 CTTACACCATAGACTTTGCCAGAACAGAGCC 3149

RESULT 14

AC013796/c 66980 bp DNA linear HTG 27-JAN-2001
 LOCUS Homo sapiens clone RP11-20H9, LOW-PASS SEQUENCE SAMPLING.

AC013796
 AC013796
 AC013796.4 GI:12583842
 HTG; HTGS PHASE0.
 Homo sapiens (human)
 Homo sapiens

REFERENCE
 1 (bases 1 to 66980)
 Birren, B., Linton, L., Nusbaum, C. and Lander, E.
 Eukaryote, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 1 (bases 1 to 66980)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
 Baldwin, J., Batra, N., Becker, R., Boguslavsky, L., Boukhgalter, B.,
 Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A.,
 Cooke, P., DeArrellano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M.,
 Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D., Gage, D.,
 Galagan, J., Gardy, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
 Howland, J. C., Johnson, R., Jones, C., Kann, L., Karas, A., Klein, J.,
 Lechoczy, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N.,
 McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrim, J.,
 Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,
 Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,
 Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
 Tesfaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,
 Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.

Direct Submission
 Submitted (15-NOV-1999) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Jan 27, 2001 this sequence version replaced gi:7582586.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Genome Center
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 Project Information
 Center project name: L3904
 Center clone name: 20_H_9

NOTE: This record contains 83 individual
 sequencing reads that have not been assembled into
 contigs. Runs of N are used to separate the reads
 and the order in which they appear is completely
 arbitrary. Low-pass sequence sampling is useful for
 identifying clones that may be gene-rich and allows
 overlap relationships among clones to be deduced.
 However, it should not be assumed that this clone
 will be sequenced to completion. In the event that
 the record is updated, the accession number will
 be preserved.

1 683: contig of 683 bp in length
 * 684 783: gap of 100 bp
 * 784 1488: contig of 705 bp in length
 * 1489 2296: gap of 100 bp
 * 1589 2296: contig of 708 bp in length
 * 2297 3101: gap of 100 bp
 * 2297 3101: contig of 705 bp in length
 * 3102 3201: gap of 100 bp
 * 3202 3916: contig of 715 bp in length
 * 3917 4016: gap of 100 bp
 * 4017 4737: contig of 721 bp in length
 * 4738 4837: gap of 100 bp
 * 4838 5557: contig of 720 bp in length
 * 5558 6372: gap of 100 bp
 * 6373 6472: gap of 100 bp
 * 6473 7155: contig of 683 bp in length
 * 7156 7255: gap of 100 bp
 * 7256 7957: contig of 702 bp in length
 * 7958 8057: gap of 100 bp
 * 8058 8773: contig of 716 bp in length
 * 8774 8873: gap of 100 bp
 * 8874 9577: contig of 704 bp in length
 * 9578 9677: gap of 100 bp
 * 9678 10392: contig of 715 bp in length
 * 10393 10492: gap of 100 bp
 * 10493 11179: contig of 687 bp in length
 * 11180 11279: gap of 100 bp
 * 11280 11984: contig of 705 bp in length
 * 11985 12084: gap of 100 bp
 * 12085 12791: contig of 707 bp in length
 * 12792 12891: gap of 100 bp
 * 12892 13600: contig of 709 bp in length
 * 13601 13700: gap of 100 bp
 * 13701 14409: contig of 709 bp in length
 * 14410 14509: gap of 100 bp
 * 14510 15221: contig of 712 bp in length
 * 15222 15321: gap of 100 bp
 * 15322 16025: contig of 704 bp in length
 * 16026 16125: gap of 100 bp
 * 16126 16822: contig of 697 bp in length
 * 16823 16922: gap of 100 bp
 * 16923 17621: contig of 699 bp in length
 * 17622 17721: gap of 100 bp
 * 17722 18423: contig of 702 bp in length
 * 18424 18523: gap of 100 bp
 * 18524 19241: contig of 718 bp in length
 * 19242 19341: gap of 100 bp
 * 19342 20056: contig of 715 bp in length
 * 20057 20156: gap of 100 bp
 * 20157 20870: contig of 714 bp in length
 * 20871 20970: gap of 100 bp

* 20971 21671: contig of 701 bp in length
 * 21672 21771: gap of 100 bp
 * 21772 22480: contig of 709 bp in length
 * 22481 22580: gap of 100 bp
 * 22581 23293: contig of 713 bp in length
 * 23294 23393: gap of 100 bp
 * 23394 24095: contig of 702 bp in length
 * 24096 24195: gap of 100 bp
 * 24196 24916: contig of 721 bp in length
 * 24917 25016: gap of 100 bp
 * 25017 25726: contig of 710 bp in length
 * 25727 25826: gap of 100 bp
 * 25827 26334: contig of 708 bp in length
 * 26335 26634: gap of 100 bp
 * 26635 27333: contig of 699 bp in length
 * 27334 27433: gap of 100 bp
 * 27434 28131: contig of 698 bp in length
 * 28132 28231: gap of 100 bp
 * 28232 28926: contig of 695 bp in length
 * 28927 29026: gap of 100 bp
 * 29027 29758: contig of 732 bp in length
 * 29759 29858: gap of 100 bp
 * 29859 30566: contig of 708 bp in length
 * 30567 30666: gap of 100 bp
 * 30667 31368: contig of 702 bp in length
 * 31369 31468: gap of 100 bp
 * 31469 32183: contig of 715 bp in length
 * 32184 32283: gap of 100 bp
 * 32284 32981: contig of 698 bp in length
 * 32982 33081: gap of 100 bp
 * 33082 33798: contig of 717 bp in length
 * 33799 33898: gap of 100 bp
 * 33899 34634: contig of 736 bp in length
 * 34635 34734: gap of 100 bp
 * 34735 35433: contig of 699 bp in length
 * 35434 35533: gap of 100 bp
 * 35534 36253: contig of 720 bp in length
 * 36254 36353: gap of 100 bp
 * 36354 37079: contig of 726 bp in length
 * 37080 37179: gap of 100 bp
 * 37180 37893: contig of 714 bp in length
 * 37894 37993: gap of 100 bp
 * 37994 38700: contig of 707 bp in length
 * 38701 38800: gap of 100 bp
 * 38801 39505: contig of 705 bp in length
 * 39506 39605: gap of 100 bp
 * 39606 40316: contig of 711 bp in length
 * 40317 40416: gap of 100 bp
 * 40417 41131: contig of 715 bp in length
 * 41132 41231: gap of 100 bp
 * 41232 41939: contig of 708 bp in length
 * 41940 42039: gap of 100 bp
 * 42040 42748: contig of 709 bp in length
 * 42749 42848: gap of 100 bp
 * 42849 43555: contig of 707 bp in length
 * 43556 43655: gap of 100 bp
 * 43656 44368: contig of 713 bp in length
 * 44369 44468: gap of 100 bp
 * 44469 45171: contig of 703 bp in length
 * 45172 45271: gap of 100 bp
 * 45272 45943: contig of 672 bp in length
 * 45944 46043: gap of 100 bp
 * 46044 46749: contig of 706 bp in length
 * 46750 46849: gap of 100 bp
 * 46850 47550: contig of 701 bp in length
 * 47551 47650: gap of 100 bp
 * 47651 48350: contig of 700 bp in length
 * 48351 48450: gap of 100 bp
 * 48451 49148: contig of 698 bp in length
 * 49149 49248: gap of 100 bp
 * 49249 49961: contig of 713 bp in length
 * 49962 50061: gap of 100 bp
 * 50062 50751: contig of 690 bp in length

* 50752 50851: gap of 100 bp
 * 50852 51567: contig of 716 bp in length
 * 51568 51667: gap of 100 bp
 * 51668 52400: contig of 733 bp in length
 * 52401 52500: gap of 100 bp
 * 52501 53225: contig of 725 bp in length
 * 53226 53325: gap of 100 bp
 * 53326 53974: contig of 649 bp in length
 * 53975 54074: gap of 100 bp
 * 54075 54781: contig of 707 bp in length
 * 54782 54881: gap of 100 bp
 * 54882 55608: contig of 727 bp in length
 * 55609 55708: gap of 100 bp
 * 55709 56426: contig of 718 bp in length
 * 56427 56526: gap of 100 bp
 * 56527 57244: contig of 718 bp in length
 * 57245 57344: gap of 100 bp
 * 57345 58051: contig of 707 bp in length
 * 58052 58151: gap of 100 bp
 * 58152 58860: contig of 709 bp in length

Query Match 45.4%; Score 108; DB 2; Length 66980;
 Best Local Similarity 96.7%; Pred. No. 4.9e-21;
 Matches 119; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 114 ATGCATCCATTGTTCTGTTTCATTTCATTCATCATCATACCCACCTGCTGAGCTTTTCAC 173
 |||||
 Db 53225 ATGCATCCATTGTTCTGTTTCATTTCATTCATCATACCCA-CTGCCNTGAGCTTTTCAC 53167
 |||||

QY 174 CTCTGACTCCCTTAACCTCCATCCATCAGACCTCTAGCACCATAAGACTCTGCAGAACTGAGAA 233
 |||||
 Db 53166 CTTGACTCCCTTAACCTCCATCCATCAGACCTTAGCACCATAAGACTCTGCAGAACTGAGAA 53107
 |||||

QY 234 GCC 236
 ||||
 Db 53106 GCC 53104

RESULT 15
 AC114854
 LOCUS
 DEFINITION Rattus norvegicus clone CH230-287P8, *** SEQUENCING IN PROGRESS
 AC114854 linear HTG 15-NOV-2002
 AC114854 Rattus norvegicus
 AC114854.4 GI:25007034
 HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ENRICHED.
 KEYWORDS Rattus norvegicus (Norway rat)
 SOURCE Rattus norvegicus
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 189486)
 Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,
 Allen, C., Allen, H., Alsebrook, S., Amin, A., Anguiano, D.,
 Anyalebechi, V., Ayodeji, A., Ayodeji, M., Baca, E., Baden, H.,
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
 Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
 Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, J.,
 Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
 Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
 Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
 Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
 Gunaratne, P., Hadland, W., Hamil, C., Hamilton, C., Hamilton, K.,
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
 Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hognes, M.,
 Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,

Karpachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuewa, L., Loulsged, H., Lozano, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartine, M., Mahmood, M., Malloy, K., Mangum, A., Mangun, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokemele, O., Okwunonu, G., Olarunpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkoch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausen, A., Weiser, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

ORIGIN

Query Match 19.9%; Score 47.4; DB 2; Length 189486;
 Best Local Similarity 63.7%; Pred. No. 0.0034;
 Matches 72; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 99 CCATCGCTTCCCTTCATGCATCCATTCCTGTTCTGTTTCATTCATTCATCCATCCACCTG 158
 DB 125370 CCACATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCC 125429

QY 159 CCTCTGAGTTTTCACCTTCGACTCCCTTAAGTTCATCCATCCATCCATCCATCCATCC 211
 DB 125430 ACTCATCCATTTCACTGTTTCATCTCTATCCATAAAGTTCATCCATCCATCCATCC 125482

Search completed: April 7, 2005, 12:50:07
 Job time : 1140.27 secs

FEATURES

source

1. 189486

/organism="Rattus norvegicus"

/mol type="genomic DNA"

/db xref="taxon:10116"

/clone="CH230-287P8"

1070..2156

/notes="clone boundary"

clone_end:Sp6

site:

end sequence:RXAHT88TV"

7765..9270

/notes="wgs contig"

21532..23046

/notes="wgs contig"

94040..94123

/notes="clone boundary"

clone_end:T7

end sequence:BZ135992"

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GSMY
 Center clone name: CH230-287P8
 ----- Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 168507 bases at least Q40
 Consensus quality: 171284 bases at least Q30
 Consensus quality: 173231 bases at least Q20
 Estimated insert size: 174384; sum-of-contigs estimation
 Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 1 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 * 1 189486: contig of 189486 bp in length.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2005, 06:13:29 ; Search time 172.895 Seconds
(without alignments)
8149.338 Million cell updates/sec

Title: US-09-977-579-4_COPY_1024_1261
Perfect score: 238
Sequence: 1 aacaggagcagtgtgacatg.....tgccagaactgagaagccgg 238

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04.*

- 1: Geneseqn1980s.*
- 2: Geneseqn1990s.*
- 3: Geneseqn2000s.*
- 4: Geneseqn2001as.*
- 5: Geneseqn2001bs.*
- 6: Geneseqn2002as.*
- 7: Geneseqn2002bs.*
- 8: Geneseqn2003as.*
- 9: Geneseqn2003bs.*
- 10: Geneseqn2003cs.*
- 11: Geneseqn2003ds.*
- 12: Geneseqn2004as.*
- 13: Geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 238 | 100.0 | 1261 | 3 | AAC67837 Human bet |
| 2 | 238 | 100.0 | 1261 | 10 | ADB78651 Human ion |
| 3 | 238 | 100.0 | 1261 | 10 | ACF57870 Human SCN |
| 4 | 236 | 99.2 | 1510 | 4 | Aaf84146 Human nov |
| 5 | 236 | 99.2 | 4052 | 6 | ABA93727 Human sig |
| 6 | 216 | 90.8 | 4625 | 4 | AAL04971 Human rep |
| 7 | 216 | 90.8 | 4625 | 4 | ABL97864 Human tes |
| 8 | 131.8 | 55.4 | 493 | 13 | ADQ55080 Novel can |
| 9 | 40.6 | 17.1 | 220756 | 12 | ADG86300 Human SMR |
| 10 | 40.6 | 17.1 | 233380 | 11 | ACN44282 Human gen |
| 11 | 38.6 | 16.2 | 11089 | 4 | AAK79653 Human imm |
| 12 | 38 | 16.0 | 26891 | 11 | ADL27167 Mouse gen |
| 13 | 38 | 16.0 | 27079 | 9 | ADA03089 Mouse lfn |
| 14 | 38 | 16.0 | 27079 | 9 | ADA66373 Mouse lfn |
| 15 | 38 | 16.0 | 27079 | 10 | ADB72827 Mouse lfn |
| 16 | 38 | 16.0 | 38360 | 6 | Abq75124 Human pl5 |
| 17 | 37.8 | 15.9 | 3768 | 6 | ABL69598 Prostate |
| 18 | 37.8 | 15.9 | 3768 | 6 | ABL63103 Breast ca |
| 19 | 37.8 | 15.9 | 3768 | 6 | ABL62894 Breast ca |
| 20 | 37.8 | 15.9 | 3768 | 6 | ABK64794 Human ben |

| | | | | | | |
|------|------|------|--------|----|------------|--------------------|
| C 21 | 37.8 | 15.9 | 3768 | 8 | ACA64938 | Human myo |
| C 22 | 37.8 | 15.9 | 3768 | 11 | ADP65339 | Human myo |
| C 23 | 37.8 | 15.9 | 8222 | 8 | ACF62757 | Human myo |
| C 24 | 37.8 | 15.9 | 8222 | 8 | ACF62815 | Colon can |
| C 25 | 37.8 | 15.9 | 8222 | 8 | ACF62793 | Colon can |
| C 26 | 37.8 | 15.9 | 11222 | 10 | ADB54189 | Pretrate |
| C 27 | 37.8 | 15.9 | 11222 | 10 | ADB54189 | Pretrate |
| C 28 | 37.8 | 15.9 | 11222 | 10 | ADB54001 | MB genomi |
| C 28 | 37.8 | 15.9 | 11222 | 10 | ADB54317 | Pretrate |
| C 29 | 37.6 | 15.8 | 1367 | 3 | AAA81673 | N. mening |
| C 30 | 37.4 | 15.7 | 1576 | 4 | AAK73497 | Human imm |
| C 31 | 37.4 | 15.7 | 7954 | 4 | AAK73498 | Human imm |
| C 32 | 37.2 | 15.6 | 110000 | 13 | ABD32911_6 | Continuation (7 of |
| C 33 | 37 | 15.5 | 2632 | 5 | AAC90602 | Rat sodiu |
| C 34 | 37 | 15.5 | 3108 | 5 | AAC90600 | Rat sodiu |
| C 35 | 37 | 15.5 | 6419 | 6 | ABL32266 | Human imm |
| C 36 | 36.6 | 15.4 | 4743 | 4 | AAI05210 | Human rep |
| C 37 | 36.6 | 15.4 | 4743 | 4 | ABL98093 | Human tes |
| C 38 | 36.6 | 15.4 | 4747 | 4 | AAI05209 | Human rep |
| C 39 | 36.6 | 15.4 | 4747 | 4 | ABL98092 | Human tes |
| C 40 | 36.4 | 15.3 | 29040 | 11 | ADL27170 | Human gen |
| C 41 | 36.4 | 15.3 | 29322 | 9 | ADA03092 | Human LFN |
| C 42 | 36.4 | 15.3 | 29322 | 9 | ADA66376 | Human LFN |
| C 43 | 36.4 | 15.3 | 29322 | 10 | ADB72830 | Human LFN |
| C 44 | 36.4 | 15.3 | 140152 | 12 | ADP03002 | Human hou |
| C 45 | 36.4 | 15.3 | 140152 | 13 | ADS88500 | Human hou |

ALIGNMENTS

RESULT 1

AAC67837
ID AAC67837 standard; cDNA; 1261 BP.

AC AAC67837;

DT 15-FEB-2001 (first entry)

XX Human beta3 CDNA.

XX Human; beta sub-unit; beta3; analgesic; anticonvulsant;
KW cerebroprotective; vasotropic; cardiac; nootropic; cytostatic;
KW dermatological; gene therapy; voltage-gated sodium channel; pain;
KW epilepsy; stroke; ischaemia; heart disease; Jacobsen Syndrome;
KW familial nonchromaffin paraganglioma; phenylketonuria;
KW Charcot Marie Tooth disease; ss.

OS Homo sapiens.

XX WO200063367-A1.

XX 26-OCT-2000.

XX 24-FEB-2000; 2000WO-EP001783.

XX 15-APR-1999; 99US-0129473P.

XX (WARN) WARNER LAMBERT CO.

XX (UYCA-) UNIV CAMBRIDGE TECH SERVICES LTD.

XX Cox P, Dixon A, Jackson A, Morgan K;

XX WPI; 2000-665241/64.

XX P-PSDB; AAB36002.

XX Novel nucleic acids encoding a beta-3 subunit from a voltage-gated sodium channel, and their corresponding polypeptides, useful for detecting and treating sodium channel-associated conditions, e.g. pain, epilepsy and stroke.

XX Claim 10; Page 70-71; 88pp; English.

XX The present sequence is given in the claims of a specification relating

CC to a novel family of beta sub-unit proteins from a voltage-gated sodium
CC channel. Human and rat beta sub-units, which have been collectively
CC identified as beta3, have been isolated. The polynucleotides and
CC polypeptides are useful for screening for agonists and antagonists of
CC sodium channels. The agonists, antagonists, proteins and nucleic acids
CC may be used diagnosing of treating diseases or conditions associated with
CC voltage-gated sodium channels, e.g. pain, epilepsy, stroke, ischaemia,
CC heart disease, Jacobson Syndrome, Familial Nonchromaffin Paraganglioma,
CC Phenylketonuria and Charcot Marie Tooth disease
XX
XX
XX Sequence 1261 BP; 263 A; 366 C; 372 G; 260 T; 0 U; 0 Other;

Query Match 100.0%; Score 238; DB 3; Length 1261;
Best Local Similarity 100.0%; Pred. No. 6.4e-69;
Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AACAGGAGCAGTGTGACATGAGTGGCTTGAACACCTGAGGACTGGACATCCCATGTTTC 60
DB 1024 AACAGGAGCAGTGTGACATGAGTGGCTTGAACACCTGAGGACTGGACATCCCATGTTTC 1083
QY 61 AGCAATGTCAATGGCATCAGGAGGGGGCCCCAAGGGCCCCATCGCTTCCCTTCATGCAATC 120
DB 1084 AGCAATGTCAATGGCATCAGGAGGGGGCCCCAAGGGCCCCATCGCTTCCCTTCATGCAATC 1143
QY 121 CATTGTTCTGTTCAATTCATTCATCCATACATCCACCTGGCTCTGAGCTTTACCTCTGAC 180
DB 1144 CATTGTTCTGTTCAATTCATTCATCCATACATCCACCTGGCTCTGAGCTTTACCTCTGAC 1203
QY 181 TCCTTAACCTCCATCAGACCTCTACGCACCATTAAGACTCTGCCAGAACTGAGAAGCGG 238
DB 1204 TCCTTAACCTCCATCAGACCTCTACGCACCATTAAGACTCTGCCAGAACTGAGAAGCGG 1261

RESULT 2
ADB78651
ID ADB78651 standard; cDNA; 1261 BP.
XX ADB78651;
XX
XX 04-DEC-2003 (first entry)
DE Human ion channel subunit cDNA mutant SCN1A exon 1 SEQ ID NO:22.
XX
XX ss; gene; mutant; ion channel; ion channel subunit; ICS; nootropic;
KW neuroprotective; inotropic; antipyrretic; antiarrhythmic; antimigraine;
KW antidepressant; antiparkinsonian; neuroleptic; tranquiliser; analgesic;
KW nephrotoxic; antidiabetic; ophthalmological; epilepsy;
KW ion channel dysfunction; human.
XX
XX Synthetic.
OS Homo sapiens.
XX WO2003008574-A1.
XX
XX 10-JAN-2003.
XX
XX 08-JUL-2002; 2002WO-AU0000910.
XX
XX 18-JUL-2001; 2001AU-00006452.
PR 05-MAR-2002; 2002AU-00000910.
PR 13-MAY-2002; 2002AU-00002292.
XX
XX (BION-) BIONOMICS LTD.
PA (WALL/) WALLACE R W.
PA
PI Mulley JC, Harkin LA, Dibbens LM, Phillips HA, Heron SE;
PI Berkovic SF, Scheffer IE;
XX
XX WPI; 2003-239332/23.
DR
XX Identifying predisposition to an ion channel dysfunction, such as
PT periodic paralysis, cardiac arrhythmias, migraine, Alzheimer's disease,
PT schizophrenia, anxiety and depression, by detecting encoding-gene

PT mutation events.
XX Claim 6; SEQ ID NO 22; 106pp; English.
PS
XX
XX The invention relates to a novel method for identifying a subject
CC predisposed to a disorder associated with ion channel dysfunction. The
CC method comprises ascertaining if at least one of the genes encoding ion
CC channel subunits (ICS) has undergone a mutation event so that a cDNA
CC derived from the subject has any of 134 nucleotide sequences. The method
CC of the invention has nootropic, neuroprotective, inotropic, antipyrretic,
CC antiarrhythmic, antimigraine, antidepressant, antiparkinsonian,
CC neuroleptic, tranquiliser, analgesic, nephrotoxic, antidiabetic, and
CC ophthalmological activity. A polynucleotide of the invention acts as an
CC ion channel agonist, or ion channel antagonist. The methods, isolated
CC nucleic acids, polypeptides, antibody, selective agonist, antagonist or
CC modulator of an ion channel, cells and genetically modified non-human
CC animal, are useful for the diagnosis and treatment of epilepsy and/or a
CC disorder associated with ion channel dysfunction, such as hyper- or hypo-
CC kalemic periodic paralysis, myotonias, malignant hyperthermia, myasthenia,
CC cardiac arrhythmias, episodic ataxia, migraine, Alzheimer's
CC disease, Parkinson's disease, schizophrenia, hyperkplexia, anxiety,
CC depression, phobic obsessive symptoms, neuropathic pain, inflammatory
CC pain, chronic/acute pain, Bartter's syndrome, polycystic kidney disease,
CC Dent's disease, hyperinsulinaemic hypoglycaemia of infancy, cystic
CC fibrosis, congenital stationary night blindness and total colour
CC blindness. The present sequence represents a mutant cDNA of the
CC invention. The sequence data for this patent is not represented in the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pat_sequences.
XX
XX Sequence 1261 BP; 263 A; 365 C; 372 G; 261 T; 0 U; 0 Other;

Query Match 100.0%; Score 238; DB 10; Length 1261;
Best Local Similarity 100.0%; Pred. No. 6.4e-69;
Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AACAGGAGCAGTGTGACATGAGTGGCTTGAACACCTGAGGACTGGACATCCCATGTTTC 60
DB 1024 AACAGGAGCAGTGTGACATGAGTGGCTTGAACACCTGAGGACTGGACATCCCATGTTTC 1083
QY 61 AGCAATGTCAATGGCATCAGGAGGGGGCCCCAAGGGCCCCATCGCTTCCCTTCATGCAATC 120
DB 1084 AGCAATGTCAATGGCATCAGGAGGGGGCCCCAAGGGCCCCATCGCTTCCCTTCATGCAATC 1143
QY 121 CATTGTTCTGTTCAATTCATTCATCCATACATCCACCTGGCTCTGAGCTTTACCTCTGAC 180
DB 1144 CATTGTTCTGTTCAATTCATTCATCCATACATCCACCTGGCTCTGAGCTTTACCTCTGAC 1203
QY 181 TCCTTAACCTCCATCAGACCTCTACGCACCATTAAGACTCTGCCAGAACTGAGAAGCGG 238
DB 1204 TCCTTAACCTCCATCAGACCTCTACGCACCATTAAGACTCTGCCAGAACTGAGAAGCGG 1261

RESULT 3
ACF57870
ID ACF57870 standard; cDNA; 1261 BP.
XX
XX ACF57870;
XX
XX 15-JAN-2004 (first entry)
XX
XX Human SCN3B protein encoding cDNA.
XX
XX SCN1A; sodium channel type 1 alpha-subunit; anticonvulsant; analgesic;
KW neuroprotective; anesthetic; cytosstatic; cerebroprotective; cardiant;
KW hypotensive; gene therapy; SCN3B; human; gene; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH 376.1023
CDS /*tag= a
FT /product= "SCN3B"
FT

XX WO2003072751-A2.
 XX 04-SEP-2003.
 XX 25-FEB-2003; 2003WO-US006010.
 XX 25-FEB-2002; 2002US-0359382P.
 XX (UYVA-) UNIV VANDERBILT.
 XX George AL, Lossin C;
 XX WPI; 2003-712725/67.
 XX P-PSDB; ABR83183.
 XX Recombinantly expressed sodium channel type 1 alpha subunit, useful in
 PT screening for modulators, for treating e.g. epilepsy.
 XX Disclosure; Page 145-147; 176pp; English.
 XX The invention relates to a recombinantly expressed and isolated human
 CC SCN1A (sodium channel type 1 alpha-subunit) (I). (I), optionally
 CC incorporated into a cell, is used to screen for specific modulators,
 CC potentially useful as anticonvulsant, antiepileptic, neuroprotective,
 CC analgesic and/or anesthetic agents, e.g. for treating severe myoclonic
 CC epilepsy of infancy, stroke, cardiac arrest, hyperkalemic paralysis,
 CC motor endplate diseases, hypertension, congestive heart failure and
 CC muscular dystrophy also to treat cancer (SCN1A is expressed in prostatic
 CC and metastatic cancer cell lines). These activities can also be provided
 CC by gene therapy vectors that express (I) or the modulators. The
 CC modulators, also antibodies directed against (I), are used to detect
 CC sodium channel polypeptides. The present sequence represents a human
 CC SCN3B protein encoding cDNA
 XX SQ Sequence 1261 BP; 263 A; 366 C; 372 G; 260 T; 0 U; 0 Other;
 Query Match 100.0%; Score 238; DB 10; Length 1261;
 Best Local Similarity 100.0%; Pred. No. 6.4e-69;
 Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AACAGGAGCAGTGTGACATGAGTGGCTGCTGAACACCTGAGGGAGCTGACATCCCATGTTTC 60
 DB 1024 AACAGGAGCAGTGTGACATGAGTGGCTGCTGAACACCTGAGGGAGCTGACATCCCATGTTTC 1083
 QY 61 AGCAATGTCAATGTCATCAGAGGGGCGCCCAAGGGGCCCCATCGCTTCCCTTCATGCATC 120
 DB 1084 AGCAATGTCAATGTCATCAGAGGGGCGCCCAAGGGGCCCCATCGCTTCCCTTCATGCATC 1143
 QY 121 CATTGTTCTGTTTCATTTCATTTCATCCATACATCCACCTGCGCTTCTGAGCTTTTCACCTCTGAC 180
 DB 1144 CATTGTTCTGTTTCATTTCATTTCATCCATACATCCACCTGCGCTTCTGAGCTTTTCACCTCTGAC 1203
 QY 181 TCCTTAACCTCCATCAGACCTCTACGACCATATAGACTCTGCCAGAACTGAGAGCGCG 238
 DB 1204 TCCTTAACCTCCATCAGACCTCTACGACCATATAGACTCTGCCAGAACTGAGAGCGCG 1261
 RESULT 4
 AAF84146
 ID AAF84146 standard; cDNA; 1510 BP.
 XX AAF84146;
 XX 07-SEP-2001 (first entry)
 XX Human novel sodium channel betal-like subunit encoding cDNA.
 XX Sodium channel; sensory neurone specific channel; betal-like subunit;
 KW SNS; therapeutic; pain; analgesic; ss.
 XX Homo sapiens.
 OS
 XX

XX Key Location/Qualifiers
 CDS 213..860
 FT /*tag= a
 FT /product= "sodium channel betal-like subunit"
 XX WO200144293-A2.
 XX 21-JUN-2001.
 XX 14-DEC-2000; 2000WO-GB004802.
 XX 17-DEC-1999; 99GB-00029970.
 XX (GLAX) GLAXO GROUP LTD.
 XX Plumpton M, Powell AJ, Sanseau P;
 XX WPI; 2001-398129/42.
 XX P-PSDB; AAB85206.
 XX Novel sub-unit for voltage-gated sodium channel proteins for producing
 PT agents useful for treating pain.
 XX Claim 4; Page 29-30; 31pp; English.
 XX The invention provides a novel betal-like sub-unit for voltage-gated
 CC sodium ion channel polypeptide, specifically a sensory neurone specific
 CC channel (SNS) subunit. The novel betal-like subunit is useful for
 CC producing a therapeutic agent which is useful treating pain in a patient.
 CC The subunit can be expressed by standard recombinant methodology. The
 CC present sequence represents a human novel sodium channel betal-like
 CC subunit encoding cDNA
 XX SQ Sequence 1510 BP; 350 A; 410 C; 414 G; 336 T; 0 U; 0 Other;
 Query Match 99.2%; Score 236; DB 4; Length 1510;
 Best Local Similarity 100.0%; Pred. No. 3.2e-68;
 Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AACAGGAGCAGTGTGACATGAGTGGCTGACACCTGAGGGAGCTGACATCCCATGTTTC 60
 DB 861 AACAGGAGCAGTGTGACATGAGTGGCTGACACCTGAGGGAGCTGACATCCCATGTTTC 920
 QY 61 AGCAATGTCAATGTCATCAGAGGGGCGCCCAAGGGGCCCCATCGCTTCCCTTCATGCATC 120
 DB 921 AGCAATGTCAATGTCATCAGAGGGGCGCCCAAGGGGCCCCATCGCTTCCCTTCATGCATC 980
 QY 121 CATTGTTCTGTTTCATTTCATTTCATCCATACATCCACCTGCGCTTCTGAGCTTTTCACCTCTGAC 180
 DB 981 CATTGTTCTGTTTCATTTCATTTCATCCATACATCCACCTGCGCTTCTGAGCTTTTCACCTCTGAC 1040
 QY 181 TCCTTAACCTCCATCAGACCTCTACGACCATATAGACTCTGCCAGAACTGAGAGGCC 236
 DB 1041 TCCTTAACCTCCATCAGACCTCTACGACCATATAGACTCTGCCAGAACTGAGAGGCC 1096
 RESULT 5
 ABA93727
 ID ABA93727 standard; cDNA; 4052 BP.
 XX ABA93727;
 XX 30-APR-2002 (first entry)
 XX Human signal transduction cDNA clone amy2_2f18.
 XX Human; foetal brain; foetal kidney; melanoma; testis; amygdala;
 KW gene therapy; ss.
 XX Homo sapiens.
 OS
 XX WO200198454-A2.
 XX

```

PD 27-DEC-2001.
XX
XX PF 25-APR-2001; 2001WO-IB002050.
XX PF 25-APR-2000; 2000US-0199380P.
XX PF (GEHU-) GERMAN HUMAN GENOME PROJECT.
XX
XX PI Wiemann S;
XX
XX DR WPI: 2002-055860/07.
XX DR P-PSDB; ABB05689.
XX
XX PT Human cDNA sequences and clones derived from human fetal brain, fetal
XX PT kidney, melanoma, testis and amygdala cDNA libraries, useful in genetic
XX PT screening and therapy.
XX
XX PS Claim 1; Page 174-175; 61lpp; English.
XX
XX CC The present invention describes assemblages and computer readable media
XX CC comprising novel human cDNA sequences and clones derived from human
XX CC foetal brain, foetal kidney, melanoma, testis and amygdala cDNA
XX CC libraries. ABA93702 to ABA93766 represent human cDNA sequences from the
XX CC present invention which encode the proteins given in ABB05662 to
XX CC ABB05729. The human cDNA sequences and clones can be used in gene
XX CC therapy. The clones may be used in a variety of applications, for example
XX CC they may be used in profiling assays, for providing large arrays of human
XX CC genetic material for implementing large-scale screening strategies and
XX CC for treating diseases via gene therapy procedures
XX
XX SQ Sequence 4052 BP; 994 A; 1033 C; 1036 G; 989 T; 0 U; 0 Other;
XX
Query Match 99.2%; Score 236; DB 6; Length 4052;
Best Local Similarity 100.0%; Pred. No. 4.8e-68;
Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AACAGGAGCAGTGTGATGAGTGGCTTGACCTGACACCTGAGGAGCTGGACATCCCATGTC 60
DB 1452 AACAGGAGCAGTGTGATGAGTGGCTTGACACCTGAGGAGCTGGACATCCCATGTC 1511
QY 61 AGCAATGTCAATGGCATCAGGAGGGGCGCCCAAGGGCCCATCGCTTCCCTTCATGCATC 120
DB 1512 AGCAATGTCAATGGCATCAGGAGGGGCGCCCAAGGGCCCATCGCTTCCCTTCATGCATC 1571
QY 121 CATTTGTTCTGTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 180
DB 1572 CATTTGTTCTGTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1631
QY 181 TCCCTAACTCCATCAGACCTTACGACCATTAAGACTCTGCCAGAACTGAGAAGCC 236
DB 1632 TCCCTAACTCCATCAGACCTTACGACCATTAAGACTCTGCCAGAACTGAGAAGCC 1687

RESULT 6
AAL04971
ID AAL04971 standard; DNA; 4625 BP.
XX
XX AC AAL04971;
XX
XX DT 21-NOV-2001 (first entry)
XX
XX DE Human reproductive system related antigen DNA SEQ ID NO: 7659.
XX
XX KW Human; reproductive system related antigen; reproductive system disorder;
XX KW cancer; gene therapy; ds.
XX
XX OS Homo sapiens.
XX
XX PN WO200155320-A2.
XX
XX PD 02-AUG-2001.
XX
XX PF 17-JAN-2001; 2001WO-US001339.
XX

```

```

XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.
XX 14-AUG-2000; 2000US-0225213P.
XX 14-AUG-2000; 2000US-0225214P.
XX 14-AUG-2000; 2000US-0225266P.
XX 14-AUG-2000; 2000US-0225267P.
XX 14-AUG-2000; 2000US-0225268P.
XX 14-AUG-2000; 2000US-0225270P.
XX 14-AUG-2000; 2000US-0225447P.
XX 14-AUG-2000; 2000US-0225757P.
XX 14-AUG-2000; 2000US-0225758P.
XX 14-AUG-2000; 2000US-0225759P.
XX 18-AUG-2000; 2000US-0226279P.
XX 22-AUG-2000; 2000US-0226881P.
XX 22-AUG-2000; 2000US-0226888P.
XX 23-AUG-2000; 2000US-0227182P.
XX 30-AUG-2000; 2000US-0227009P.
XX 30-AUG-2000; 2000US-0228924P.
XX 01-SEP-2000; 2000US-0229287P.
XX 01-SEP-2000; 2000US-0229343P.
XX 01-SEP-2000; 2000US-0229344P.
XX 01-SEP-2000; 2000US-0229345P.
XX 05-SEP-2000; 2000US-0229509P.
XX 05-SEP-2000; 2000US-0229513P.
XX 06-SEP-2000; 2000US-0230437P.
XX 06-SEP-2000; 2000US-0230438P.
XX 08-SEP-2000; 2000US-0231242P.
XX 08-SEP-2000; 2000US-0231243P.
XX 08-SEP-2000; 2000US-0231244P.
XX 08-SEP-2000; 2000US-0231413P.
XX 08-SEP-2000; 2000US-0231414P.
XX 08-SEP-2000; 2000US-0232080P.
XX 08-SEP-2000; 2000US-0232081P.
XX 12-SEP-2000; 2000US-0231968P.
XX 14-SEP-2000; 2000US-0232397P.
XX 14-SEP-2000; 2000US-0232398P.
XX 14-SEP-2000; 2000US-0232399P.
XX 14-SEP-2000; 2000US-0232400P.
XX 14-SEP-2000; 2000US-0232401P.
XX 14-SEP-2000; 2000US-0233063P.
XX 14-SEP-2000; 2000US-0233064P.
XX 14-SEP-2000; 2000US-0233065P.
XX 21-SEP-2000; 2000US-0234223P.
XX 21-SEP-2000; 2000US-0234274P.
XX 25-SEP-2000; 2000US-0234997P.
XX 25-SEP-2000; 2000US-0234998P.
XX 26-SEP-2000; 2000US-023484P.
XX 27-SEP-2000; 2000US-0235834P.
XX 27-SEP-2000; 2000US-0235835P.
XX 29-SEP-2000; 2000US-0236327P.
XX 29-SEP-2000; 2000US-0236367P.
XX 29-SEP-2000; 2000US-0236368P.
XX 29-SEP-2000; 2000US-0236369P.
XX 29-SEP-2000; 2000US-0236370P.
XX

```

| | | |
|----------|-----|---|
| XX | | Disclosure; SEQ ID NO 7659; 1297pp + Sequence Listing; English. |
| XX | | The present invention provides the protein and coding sequences of a |
| CC | | number of human reproductive system related antigens. These can be used |
| CC | | in the prevention and treatment of reproductive system disorders, |
| CC | | including cancer. The present sequence is a genomic sequence encoding a |
| CC | | protein of the invention |
| XX | | |
| SQ | | Sequence 4625 BP; 1222 A; 1130 C; 906 G; 1367 T; 0 U; 0 Other; |
| | | |
| | | Query Match 90.8%; Score 216; DB 4; Length 4625; |
| | | Best Local Similarity 100.0%; Pred.No.2.5e-61; |
| | | Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0; |
| | | |
| QY | 21 | AGGTGGCCTGAACACACTGAGGGAGCTGGACATCCCATCGTTTCAGCAATGTCATGGCATCAG 80 |
| DB | 33 | AGTGGCCTGAACACACTGAGGGAGCTGGACATCCCATCGTTTCAGCAATGTCATGGCATCAG 92 |
| | | |
| QY | 81 | GAGGGGCCCCAAAGGGCCCCATCGCTTCCTTCATGCATCATTTGTTCTGTTTCATTTCATT 140 |
| DB | 93 | GAGGGGCCCCAAAGGGCCCCATCGCTTCCTTCATGCATCATTTGTTCTGTTTCATTTCATT 152 |
| | | |
| QY | 141 | CATCCATACATCCACCTGCGCTTGAGCTTTCACCTCTGACTCCCTTAACCTCCATCAGACCT 200 |
| DB | 153 | CATCCATACATCCACCTGCGCTTGAGCTTTCACCTCTGACTCCCTTAACCTCCATCAGACCT 212 |
| | | |
| QY | 201 | CTACGGCACCATAGAAGCTCTGCCAAGAACTGAGAAGCC 236 |
| DB | 213 | CTACGGCACCATAGAAGCTCTGCCAAGAACTGAGAAGCC 248 |
| | | |
| RESULT 7 | | |
| ABL97864 | | ID ABL97864 standard; DNA; 4625 BP. |
| XX | | ABL97864; |
| XX | | 21-JUN-2002 (first entry) |
| XX | | Human testicular antigen encoding DNA fragment SEQ ID NO: 2516. |
| XX | | Human; testicular antigen; testes; cancer; metastasis; immune disorder; |
| KW | | reproductive system disorder; urinary system disorder; gene therapy; |
| KW | | cardiovascular disorder; respiratory disorder; neurological disorder; |
| KW | | gastrointestinal disease; infection; cytostatic; gene; ds. |
| OS | | Homo sapiens. |
| XX | | |
| PN | | WO200155317-A2. |
| XX | | |
| PD | | 02-AUG-2001. |
| XX | | |
| PF | | 17-JAN-2001; 2001WO-US001329. |
| XX | | |
| XX | | 31-JAN-2000; 2000US-0179065P. |
| PR | | 04-FEB-2000; 2000US-0180628P. |
| PR | | 24-FEB-2000; 2000US-0184564P. |
| PR | | 02-MAR-2000; 2000US-0186350P. |
| PR | | 16-MAR-2000; 2000US-0189874P. |
| PR | | 17-MAR-2000; 2000US-0190076P. |
| PR | | 18-APR-2000; 2000US-0198123P. |
| PR | | 19-MAY-2000; 2000US-0205515P. |
| PR | | 07-JUN-2000; 2000US-0209467P. |
| PR | | 28-JUN-2000; 2000US-0214886P. |
| PR | | 30-JUN-2000; 2000US-0215135P. |
| PR | | 07-JUL-2000; 2000US-0216647P. |
| PR | | 07-JUL-2000; 2000US-0216880P. |
| PR | | 11-JUL-2000; 2000US-0217487P. |
| PR | | 11-JUL-2000; 2000US-0217496P. |
| PR | | 14-JUL-2000; 2000US-0218290P. |
| PR | | 26-JUL-2000; 2000US-0220963P. |
| PR | | 26-JUL-2000; 2000US-0220964P. |

Db 33 AGGTGGCTGAACACCTGAGGAGCTGGACATCCCATGTTTCAGCAATGCAATGCATCAG 92
QY 81 GAGGGCGCCCAAGGGCCCATCGCTTCCTTCATGTCATCCATGTTCTGTTTCATTTCATT 140
Db 93 GAGGGCGCCCAAGGGCCCATCGCTTCCTTCATGTCATCCATGTTCTGTTTCATTTCATT 152
QY 141 CATCCATACATCCACCTGCGCTCTGAGCTTTTCACCTCTGACTCCCTAACTCCATCAGACCT 200
Db 153 CATCCATACATCCACCTGCGCTCTGAGCTTTTCACCTCTGACTCCCTAACTCCATCAGACCT 212
QY 201 CTAGGCACCATAGACTCTGCGAGAACTGAGAGCC 236
Db 213 CTAGGCACCATAGACTCTGCGAGAACTGAGAGCC 248

RESULT 8

ADQ55080
ID ADQ55080 standard; DNA; 493 BP.
XX AC ADQ55080;
XX DT 21-OCT-2004 (first entry)
XX DE Novel canine microarray-related DNA sequence SeqID6382.
XX KW canine microarray; drug screening; toxicity assay;
KW environmental pollutant; cellular response; gene expression profile;
KW toxic response; liver necrosis; fatty liver disease;
KW protein adduct formation; hepatitis; dog; ds.
XX OS Canis familiaris.
XX PN WO2004063324-A2.
XX PD 29-JUL-2004.
XX PF 05-MAY-2003; 2003WO-US013853.
XX PR 03-MAY-2002; 2002US-0377240P.
XX PA (GENE-) GENE LOGIC INC.
PA (PFIZ) PFIZER PROD INC.
XX PI Diggins JC, Porter M, Wei T;
XX WPI; 2004-561890/54.
XX DR New isolated nucleic acid molecule, useful for drug screening and
XX PT toxicity assays or for assessing the impact, including toxicity, of a
PT compound, pharmaceutical agent or environmental pollutant on a cell or
PT living organism.
XX PS Claim 1; SEQ ID NO 6382; 41pp; English.

XX This invention is related to a novel isolated canine nucleic acid
CC sequences and the construction of canine microarrays containing a
CC significant portion of the canine genome. The isolated canine nucleic
CC acid sequences of the invention may be useful for drug screening and
CC toxicity assays. The invention is therefore useful for assessing the
CC impact, including toxicity, of a compound, pharmaceutical agent or
CC environmental pollutant on a cell or living organism. The methods are
CC useful for detecting genes that are up- or down-regulated in canines in a
CC disease state. The sequences are useful as diagnostic agents or markers
CC to detect a cellular response in a sample individually or as part of a
CC gene expression profile. It is also useful as a target for agents that
CC modulate gene expression or activity. The database is useful for
CC producing electronic Northern blots that allow the user to determine the cell
CC type or tissue in which a given gene is expressed and to allow
CC determination of the abundance or expression level of a given gene in a
CC particular tissue or cell. The methods are useful for determining the
CC similarity of a toxic response to one or more individual compounds. The
CC methods are useful for predicting at least one toxic response or the
CC likelihood that a compound or test agent will induce various specific

CC pathologies such as those of the liver (liver necrosis, fatty liver
CC disease, protein adduct formation or hepatitis), those of the kidney,
CC heart, brain or testes, or other pathologies associated with at least one
CC of the toxins. The methods are also useful for predicting or elucidating
CC the potential cellular pathways influenced, induced or modulated by the
CC compound or test agent due to the similarity of the expression profile
CC compared to the profile induced by a known toxin. The present sequence is
CC that of a canine DNA sequence which was claimed for use during the
CC production of a canine microarray of the invention.

XX SQ Sequence 493 BP; 102 A; 149 C; 111 G; 131 T; 0 U; 0 Other;

Query Match 55.4%; Score 131.8; DB 13; Length 493;
Best Local Similarity 82.3%; Pred. No. 1.5e-33;
Matches 177; Conservative 0; Mismatches 32; Indels 6; Gaps 2;

QY 21 AGGTGGCTGAACACCTGAGGAGCTGGACATCCCATGTTTCAGCAATGCAATGCATCAG 80
Db 114 AGGTGGCTGAACA-CTGAAGGACTGAATATCCAGGTTTCAGTGTCTCAACGACGAG 172
QY 81 GAGGGCGCCCAAGGGCCCATCGCTTCCTTCATGTCATCCATGTTCTGTTTCATTTCATT 140
Db 173 GAGGGTGGCCCAAGGGCCCATCGCTTCCTTCATGTCATCCATGTTCTGTTTCATTTCATT 231
QY 141 CATCCATACATCCACCTGCGCTCTGAGCTTTTCACCTCTGACTCCCTAACTCCATCAGACCT 200
Db 232 ----CATGCATCGCTCACCTCTGAGCTTTTCACCTCTGACTCTGACTCTGACTCTGACTCT 287
QY 201 CTAGGCACCATAGACTCTGCGAGAACTGAGAGCC 235
Db 288 CTACACACCAAGAGACTTTGCCAGAACTGAGGAGC 322

RESULT 9

ADG86300
ID ADG86300 standard; DNA; 220756 BP.
XX AC ADG86300;
XX DT 11-MAR-2004 (first entry)
XX DE Human SMRT partial genomic DNA sequence SEQ ID NO:14.
XX KW SMRT; silencing mediator for retinoid and thyroid hormone action;
KW SMRT inhibitor; cytostatic; antiinflammatory; antiarthritic;
KW antirheumatic; antitense therapy; inflammatory disorder;
KW rheumatoid arthritis; hyperproliferative disorder; cancer; leukaemia;
KW breast cancer; human; gene; ds.
XX OS Homo sapiens.
XX PN WO2003106645-A2.
XX PD 24-DEC-2003.
XX PF 17-JUN-2003; 2003WO-US018923.
XX PR 17-JUN-2002; 2002US-00174014.
XX PA (ISIS-) ISIS PHARM INC.
XX PI Bennett CF, Freier SM, Dobie KW;
XX WPI; 2004-082184/08.
XX DR GENBANK; NT_009459.
XX PT Novel antisense compound targeted to nucleic acid encoding SMRT
PT (silencing mediator for retinoid and thyroid hormone action), useful for
PT treating animal having disease associated with SMRT such as cancer,
PT rheumatoid arthritis.
XX PS Example 15; SEQ ID NO 14; 260pp; English.

The present invention describes a compound (I) 8-50 nucleobases in length targeted to a nucleic acid molecule encoding SMRT (silencing mediator for retinoid and thyroid hormone action), where (I) specifically hybridises with the nucleic acid molecule encoding SMRT and inhibits expression of SMRT. (I) specifically hybridises with at least 8-nucleobase portion of a preferred target region on nucleic acid molecule encoding SMRT. Also described is a composition (II) comprising (I) and a carrier or diluent. (I) and (II) have cytostatic, antiinflammatory, antiarthritic and antirheumatic activities, (I) can be used in antineoplastic therapy, and as SMRT expression inhibitors. (I) is useful for inhibiting the expression of SMRT in cells or tissues. (I) is also useful for treating an animal having a disease or condition associated with SMRT, e.g., inflammatory disorder such as rheumatoid arthritis; or a hyperproliferative disorder such as cancer caused from leukaemia and breast cancer, by inhibiting the expression of SMRT. (I) is useful for diagnostics, therapeutics, prophylaxis and as research reagents and kits. The present sequence represents a partial genomic DNA sequence of human SMRT, which is used in an example from the present invention. N.B. The present sequence is designated as SEQ ID NO:12 in example 15 but corresponds to SEQ ID NO:14 in the Sequence Listing.

[illegible]

| | |
|---|--|
| RESULT 10 | |
| ACN44282 | |
| ID ACN44282 standard; DNA; 233380 BP. | |
| XX | |
| XX AC ACN44282; | |
| XX | |
| XX 18-NOV-2004 (first entry) | |
| DT | |
| XX | |
| XX Human genomic sequence hCG25303. | |
| DE | |
| XX | |
| XX Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss. | |
| KW | |
| XX Homo sapiens. | |
| XX | |
| XX WO2003073826-A2. | |
| PN | |
| XX | |
| XX 12-SEP-2003. | |
| PD | |
| XX | |
| XX 28-FEB-2003; 2003WO-US006235. | |
| PF | |
| XX | |
| XX 01-MAR-2002; 2002US-00087192. | |
| PR | |
| XX | |
| XX (SAGR-) SAGRES DISCOVERY. | |
| XX | |
| XX Morris DW; | |
| PI | |
| XX | |
| XX WPI; 2003-328604/31. | |
| DR | |
| XX | |
| XX Recombinant nucleic acid useful for diagnosis and treatment of carcinoma | |
| PT | |
| PT comprises a nucleotide sequence. | |
| XX | |
| XX Claim 1; SEQ ID NO 652; Opp; English. | |
| PS | |

The present invention relates to novel DNA and protein sequences which are associated with carcinomas. The sequences are useful for: (i) for screening drug candidates; (ii) for screening of bioactive agent capable of binding to Carcinoma Associated Protein (CAP); (iii) for screening of a bioactive agent capable of modulating the activity of CAP; (iv) for evaluating the effect of a candidate carcinoma drug; (v) for diagnosing carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip; (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for determining Carcinoma Associated (CA) gene copy number. In addition, the CA genes are useful as DNA vaccines and the CAP are useful as markers of carcinoma including lymphoma. The present sequence is one such CA coding sequence. Note: This patent is an equivalent to basic patent US2002182586A1, for which no sequence data was published

Sequence 233380 BP: 44357 A; 63089 C; 67702 G; 51928 T; 0 U 6304 Other:

| Seq | Sequence | 233380 BP; | 44357 A; | 63089 C; | 67702 G; | 51928 T; | 0 U; | 6304 Other; |
|-----------------------|----------|--|----------|------------|----------|----------|---------|-------------|
| Query Match | 17.1%; | Score | 40.6; | DB | 11; | Length | 233380; | |
| Best Local Similarity | 57.5%; | Pred. No. | 0.055; | | | | | |
| Matches | 73; | Conservative | 0; | Mismatches | 54; | Indels | 0; | Gaps |
| Qy | 46 | GGACATCCCATGTTTCAGCAATGTCATGCGAGGGGCCCCCAAGGGCCCCCATCGC | 105 | | | | | |
| Db | 55725 | GGAGCTCCCATGCTGGGAGAGATTCAATTCATTGGGAGTGTCCTCCATGCTGGGAGATCC | 55784 | | | | | |
| Qy | 106 | TTCCCTTCATGCAATCAATTTGTTCTGTTCATTCATTCATCCATCACCTGCCTCTGA | 165 | | | | | |
| Db | 55785 | ATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCTGGCAATA | 55844 | | | | | |
| Qy | 166 | GCTTTCA | 172 | | | | | |
| Db | 55845 | CTTATCA | 55851 | | | | | |

| | |
|-----------|---|
| RESULT 11 | |
| AAK79653 | |
| ID | AAK79653 standard; DNA; 11089 BP. |
| XX | |
| XX | |
| AC | AAK79653; |
| XX | |
| XX | |
| DT | 07-NOV-2001 (first entry) |
| XX | |
| DE | Human immune/haematopoietic antigen genomic sequence SEQ ID NO:34465. |
| XX | |
| XX | |
| KW | Human; immune; haematopoietic; immune/haematopoietic antigen; cancer; |
| KW | cytostatic; gene therapy; vaccine; metastasis; ds. |
| XX | |
| OS | Homo sapiens. |
| XX | |
| XX | |
| PN | WO2001571182-A2. |
| XX | |
| XX | |
| PD | 09-AUG-2001. |
| XX | |
| PF | 17-JAN-2001; 2001WO-US001354. |
| XX | |
| XX | |
| PR | 31-JAN-2000; 2000US-0179065P. |
| PR | 04-FEB-2000; 2000US-0180628P. |
| PR | 24-FEB-2000; 2000US-0184664P. |
| PR | 02-MAR-2000; 2000US-0186350P. |
| PR | 16-MAR-2000; 2000US-0189874P. |
| PR | 17-MAR-2000; 2000US-0190076P. |
| PR | 18-APR-2000; 2000US-0198123P. |
| PR | 19-MAY-2000; 2000US-0205515P. |
| PR | 07-JUN-2000; 2000US-0209467P. |
| PR | 28-JUN-2000; 2000US-0214886P. |
| PR | 30-JUN-2000; 2000US-0215135P. |
| PR | 07-JUL-2000; 2000US-0216647P. |
| PR | 07-JUL-2000; 2000US-0216880P. |
| PR | 11-JUL-2000; 2000US-0217487P. |
| PR | 11-JUL-2000; 2000US-0217456P. |
| PR | 14-JUL-2000; 2000US-0218290P. |
| PR | 26-JUL-2000; 2000US-0220963P. |
| PR | 26-JUL-2000; 2000US-0220964P. |

PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226688P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 08-SEP-2000; 2000US-0229513P.
 PR 08-SEP-2000; 2000US-0230437P.
 PR 08-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 25-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 01-NOV-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.

PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 (HUMA-) HUMAN GENOME SCI INC.
 Rosen CA, Barash SC, Ruben SM;
 WPI; 2001-483426/52.
 Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 useful for preventing, diagnosing and/or treating cancers and metastasis.
 Disclosure; SEQ ID NO 34465; 3071pp + Sequence Listing; English.
 AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
 activity, and can be used in gene therapy and vaccine production. (I)
 proteins and polynucleotides may be used in the prevention, diagnosis and
 treatment of diseases associated with inappropriate (I) expression. For
 example, they may be used to treat disorders associated with decreased
 expression by rectifying mutations or deletions in a patient's genome
 that affect the activity of (I) by expressing inactive proteins or to
 supplement the patient's own production of (I). Additionally, (I)
 polynucleotides may be used to produce the secreted (I), by inserting the
 nucleic acids into a host cell and culturing the cell to express the
 protein. (I) proteins and polynucleotides may be used to prevent,
 diagnose and treat immune/hematopoietic-related diseases, especially
 cancers and cancer metastases of hematopoietic-derived cells. AAK64703
 to AAK87694 represent human immune/hematopoietic antigen genomic
 sequences from the present invention. AAK54942 to AAK54950 and AAK82169

Query Match 16.0%; Score 38; DB 9; Length 27079;
Best Local Similarity 57.6%; Pred. No. 0.17;
Matches 68; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

| | | | |
|------|----|---|------|
| 104 | QY | GCTTCCTTCATGCATCCATGTTCTTGTTTCATTCATTCATCCATGATCCACCTGCCTCT | 163 |
| 3703 | Db | GCTTCCTCCATGCATCTATCATCTTTATCCATCCATTCATGATTCATCCATCCATCCAT | 3762 |
| 164 | QY | GAGCTTTTCACTCTGACTCCCTAACTCCCATCAGACCTTACGCACCATAAGACTCTGC | 221 |
| 3763 | Db | CCAACTATCCATCTGTGCTCCGCCACCCACCCACATCATCTACACATGCTCCATGCAGC | 3820 |

RESULT 15
ADB72827
ID ADB72827 standard; DNA; 27079 BP.

| | |
|----|---------------------------|
| AA | |
| AC | ADB72827; |
| XX | |
| DT | 04-DEC-2003 (first entry) |
| XX | |
| DE | Mouse Lfng gene. |

| | |
|----|---|
| XX | mouse; ds; cytostatic; gene therapy; vaccine; carcinoma; lymphomas; |
| KW | cancer; neoplasm; adenocarcinoma; sarcoma; gene. |
| XX | |
| XX | Mus sp. |
| XX | |
| XX | WO2003008583-A2. |
| XX | |
| XX | 30-JAN-2003. |
| XX | |
| XX | |

XX
PF 28-DEC-2001; 2001MO-03031291.

FK 02-PTA-2001; 2001US-00730380
PR 23-OCT-2001; 2001US-00004113.

FK 08-NOV-2001; 2001US-0003248Z.
PR 30-NOV-2001; 2001US-0099772Z.

XX
XX

SECRET - NOT FOR RELEASE TO THE PUBLIC

100-350-07 YJ

XX
XXXX
FI MOULTIS DW, Eingetragt BN,XX
XX
07/156667-0007;TFM DR

PS Claim 1; SEQ ID NO 655; 2304pp; English.

CC The invention relates to a novel recombination

XX Sequence 27079 BP; 5965 A; 7344 C; 7585 G; 6185 T; 0 U; 0 Other;
SQ

Db 3763 CCAACATCCCATCTGTCCCCCACCACCCACATCATCTACACATCTCCCATGCAGC 3820

Search completed: April 7, 2005, 11:09:40
Job time : 177.885 secs

| Result No. | Score | Query Match | Length | DB ID | Description | |
|------------|-------|-------------|--------|-------|----------------------|--------------------|
| | | | | | | |
| C 1 | 38.6 | 16.2 | 6242 | 4 | US-09-949-016-16490 | Sequence 16490, A |
| C 2 | 37.8 | 15.9 | 14566 | 4 | US-09-949-016-16765 | Sequence 16765, A |
| C 3 | 37.2 | 15.6 | 23640 | 4 | US-09-949-016-13646 | Sequence 13646, A |
| C 4 | 37 | 15.5 | 767677 | 4 | US-09-949-016-12147 | Sequence 12147, A |
| C 5 | 37 | 15.5 | 767677 | 4 | US-09-949-016-17361 | Sequence 17361, A |
| C 6 | 36.2 | 15.2 | 7218 | 1 | US-08-232-463-14 | Sequence 14, Appl |
| C 7 | 36.2 | 15.2 | 60276 | 4 | US-09-949-016-15004 | Sequence 15004, A |
| C 8 | 36.2 | 15.2 | 60338 | 4 | US-09-949-016-15694 | Sequence 15694, A |
| C 9 | 36 | 15.1 | 323820 | 4 | US-09-949-016-14139 | Sequence 14139, A |
| C 10 | 35.8 | 15.0 | 71278 | 4 | US-09-949-016-11851 | Sequence 11851, A |
| C 11 | 35.8 | 15.0 | 71278 | 4 | US-09-949-016-17563 | Sequence 17563, A |
| C 12 | 35.8 | 15.0 | 168174 | 4 | US-10-071-411A-63 | Sequence 63, Appl |
| C 13 | 35.8 | 15.0 | 168273 | 4 | US-10-071-411A-2 | Sequence 2, Appl |
| C 14 | 35.4 | 14.9 | 601 | 4 | US-09-949-016-74297 | Sequence 74297, A |
| C 15 | 35.4 | 14.9 | 601 | 4 | US-09-949-016-74298 | Sequence 74298, A |
| C 16 | 35.4 | 14.9 | 30656 | 4 | US-09-949-016-14613 | Sequence 14613, A |
| C 17 | 35.4 | 14.9 | 84761 | 4 | US-09-949-016-11919 | Sequence 11919, A |
| C 18 | 35.4 | 14.9 | 84763 | 4 | US-09-949-016-13914 | Sequence 13914, A |
| C 19 | 35.2 | 14.8 | 18900 | 4 | US-09-949-016-16889 | Sequence 16889, A |
| C 20 | 35.2 | 14.8 | 30000 | 4 | US-10-007-010-10 | Sequence 10, Appl |
| C 21 | 35.2 | 14.8 | 53562 | 4 | US-09-949-016-16286 | Sequence 16286, A |
| C 22 | 34.4 | 14.5 | 601 | 4 | US-09-949-016-16286 | Sequence 16286, A |
| C 23 | 34.4 | 14.5 | 601 | 4 | US-09-949-016-79198 | Sequence 79198, A |
| C 24 | 34.4 | 14.5 | 601 | 4 | US-09-949-016-91783 | Sequence 91783, A |
| C 25 | 34.4 | 14.5 | 601 | 4 | US-09-949-016-91783 | Sequence 91784, A |
| C 26 | 34.4 | 14.5 | 601 | 4 | US-09-949-016-91785 | Sequence 91785, A |
| C 27 | 34.4 | 14.5 | 601 | 4 | US-09-949-016-91786 | Sequence 91786, A |
| C 28 | 34.4 | 14.5 | 601 | 4 | US-09-949-016-194724 | Sequence 194724, A |

```
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16765
; LENGTH: 14566
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16765

Query Match
Best Local Similarity 15.9%; Score 37.8; DB 4; Length 14566;
Matches 63; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 105 CTTCCCTTCATCATCATGTTCTGTTTCATTCATCCATCCATCCACCTGCTCTG 164
Db 528 CATCCCTCCATCATCATCCACCCATCCCTCATTCATCCATCATTCCTCCATTC 469
QY 165 AGTTTCACTCTGACTCCCTCACTCACTCCATCCATCCAGACCTCTACGCACC 209
Db 468 ATCATCCCTCCCTCCATCCATCTATCCATCCATCCACCCACCATC 424

RESULT 3
US-09-949-016-13646
; Sequence 13646, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13646
; LENGTH: 23640
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(23640)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13646

Query Match
Best Local Similarity 15.6%; Score 37.2; DB 4; Length 23640;
Matches 54; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 107 TCCCTTCATGCATCCATGTTCTGTTTCATTCATCCATCCATCCACCTGCTCTGAG 166
Db 7218 TCCATCCATGCACCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCC 7277
QY 167 CTTTCACTCTGACTCCCTCAAC 189
Db 7278 CTTTACTTCTACTACCCATC 7299

RESULT 4
US-09-949-016-12147/c
; Sequence 12147, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12147
; LENGTH: 76767
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(76767)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12147

Query Match
Best Local Similarity 15.5%; Score 37; DB 4; Length 76767;
Matches 61; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 109 CTTTCATGCATCCATGTTCTGTTTCATTCATCCATCCATCCACCTGCTCTGAGCT 168
Db 542579 CTTTCCTCCATCCATCCATCCCTTCCATCCATCCACCCACATCCATCCATCTTCT 542520
QY 169 TTCACCTCTGACTCCCTCACTCCATCCATCCAGACCTCTACGCACC 209
Db 542519 TACATCCATCCACCCATCCATCCCTTCTTCATCCATCCATC 542479

RESULT 5
US-09-949-016-17361/c
; Sequence 17361, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17361
; LENGTH: 76767
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(76767)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17361

Query Match
Best Local Similarity 15.5%; Score 37; DB 4; Length 76767;
Matches 61; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 109 CTTTCATGCATCCATGTTCTGTTTCATTCATCCATCCATCCACCTGCTCTGAGCT 168
Db 542579 CTTTCCTCCATCCATCCATCCCTTCCATCCATCCACCCACATCCATCCATCTTCT 542520
QY 169 TTCACCTCTGACTCCCTCACTCCATCCATCCAGACCTCTACGCACC 209
Db 542519 TACATCCATCCACCCATCCATCCCTTCTTCATCCATCCATC 542479
```


Qy 169 TTCACTCTGACTCCCTTAAGTCCATCCAGACTCTTACGCACC 209
Db 542519 TACATCCATCCACCACTCCATCTTCTTATCCATCCATC 542479

RESULT 6

US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOVLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F1s
US-08-232-463-14

Query Match 15.2%; Score 36.2; DB 1; Length 7218;
Best Local Similarity 8.4%; Pred.No.0.068; 77; Indels 0; Gaps 0;
Matches 17; Conservative 109; Mismatches 77; Indels 0; Gaps 0;
Qy 20 GAGGTGGCGTGAACACCTGAGGGAGCTGACATCCATCCATGTTTCAGCAATGTCATGGCATCA 79
Db 1035 GAGCTTGCTGCAGTGCAGGGAGCTTGCATATTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1094
Qy 80 GGAGGGCGCCCAAGGGCCCATCGCTTCCCTTCATGCATCCATGTTCTGTTCTTCATTCAT 139
Db 1095 YY 1154
Qy 140 TCATCCATACATCCACCTGCTTTCAGCTTTCACCTCTGACTCTGACTCCCTAATCCATCACC 199
Db 1155 YY 1214
Qy 200 TCTACGCACCAATAGACTCTGCC 222

Db 1215 YYYYYYYYYYYYYYYYYYYY 1237

RESULT 7

US-09-949-016-15004
; Sequence 15004, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15004
; LENGTH: 60276
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(60276)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15004

Query Match 15.2%; Score 36.2; DB 4; Length 60276;
Best Local Similarity 57.5%; Pred.No.0.21; 48; Indels 0; Gaps 0;
Matches 65; Conservative 0; Mismatches 48; Indels 0; Gaps 0;
Qy 111 TTTCATCATCCATGTTTCTGTTTCATTCATTCATCCATCCATCCATCCCTGCTCTGAGCTTT 170
Db 11773 TCCATCGATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCAT 11832
Qy 171 CACTCTGACTCCCTTAAGTCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCC 223
Db 11833 TCACCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCC 11885

RESULT 8

US-09-949-016-15694
; Sequence 15694, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15694
; LENGTH: 60338
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(60338)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15694

```
Query Match      15.2%; Score 36.2; DB 4; Length 60338;
Best Local Similarity 57.5%; Pred. No. 0.21; Indels 0; Gaps 0;
Matches 65; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 111 TTTCATCATCCATTTGTTCTGTTTCATTCATTCATCCATACATCCATCGCTGCTGAGCTTT 170
    |||||
Db 11835 TCATCATCCATCCATCCATCCATCCATTCAGCCATCCATCCATCCATCCATCCATCCAT 11894
    |||||

QY 171 CACCTCTGACTCCCTTAACCTCACTCAGACCTCTACGCACCATAGACTCTGCGCA 223
    |||||
Db 11895 TCACCCATCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCAT 11947

RESULT 9
US-09-949-016-14139/c
; Sequence 14139, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14139
; LENGTH: 323820
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(323820)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14139

Query Match      15.1%; Score 36; DB 4; Length 323820;
Best Local Similarity 70.6%; Pred. No. 0.58; Indels 20; Gaps 0;
Matches 48; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

QY 89 CCCAAGGCCCGCCATCGCTTCCCTTCATGCATCCATTTGTTCTGTTTCATTCATCCATC 148
    |||||
Db 88457 CTCAAGGCCCGCTTCTCTCCATCCATCCATCCATCCATCCATCCATCCATCCATCC 88398
    |||||

QY 149 CATCCACC 156
    |||||
Db 88397 CATCCAAC 88390

RESULT 10
US-09-949-016-11851/c
; Sequence 11851, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
```

```
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11851
; LENGTH: 71278
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(71278)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-11851

Query Match      15.0%; Score 35.8; DB 4; Length 71278;
Best Local Similarity 56.3%; Pred. No. 0.31; Indels 52; Gaps 0;
Matches 67; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 105 CTTCCCTTCATGCATCCATTTGTTCTGTTTCATTCATTCATCCATACATCCATCGCTCTG 164
    |||||
Db 40115 CATCCATCATCTATCCATCCATCTCTGTCTATATTCATCCATCCATCCATCCATTC 40056
    |||||

QY 165 AGCTTTACCTCTGACTCCCTTAACCTCACTCAGACCTCTACGCACCATAGACTCTGCGCA 223
    |||||
Db 40055 ACCATCCACCCATCAATCCATCCATCCATCCATCCATCCATCCATCCATCCATCTCT 39997
    |||||

RESULT 11
US-09-949-016-17563/c
; Sequence 17563, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17563
; LENGTH: 71278
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(71278)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17563

Query Match      15.0%; Score 35.8; DB 4; Length 71278;
Best Local Similarity 56.3%; Pred. No. 0.31; Indels 52; Gaps 0;
Matches 67; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 105 CTTCCCTTCATGCATCCATTTGTTCTGTTTCATTCATTCATCCATACATCCATCGCTCTG 164
    |||||
Db 40115 CATCCATCATCTATCCATCCATCTCTGTCTATATTCATCCATCCATCCATCCATTC 40056
    |||||

QY 165 AGCTTTACCTCTGACTCCCTTAACCTCACTCAGACCTCTACGCACCATAGACTCTGCGCA 223
    |||||
Db 40055 ACCATCCACCCATCAATCCATCCATCCATCCATCCATCCATCCATCCATCTCTCTATCTGTCA 39997
    |||||

RESULT 12
US-10-071-411A-63
; Sequence 63, Application US/10071411A
; Patent No. 6797475
; GENERAL INFORMATION:
; APPLICANT: Glenn Barnes
```

| | Query Match | 15.0% | Score 35.8; | DB 4; | Length 168273; |
|----|-----------------------|---|-----------------|-----------|----------------|
| | Best Local Similarity | 56.3%; | Pred. No. 0.48; | | |
| | Matches 67; | Conservative 0; | Mismatches 52; | Indels 0; | Gaps 0; |
| Qy | 105 | CTTCCCTTCATGCATCCATTTGTTCTGTCATTCAATCCACATACATCCACCTGCCTCTG | 164 | | |
| Db | 12256 | CATCCATCCATCTATCCATCCATCTCTGTCTATCTTCCATTCATCCACCTCATTC | 12315 | | |
| Qy | 165 | AGCTTTTCACCTCTGACTTCCCTAACTCCATCAGACCTCTTACGACCACCATAGACTCTGCCA | 223 | | |

RESULT 15
US-09-949-016-74298
Sequence 74298, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: Fast-SEQ for Windows Version 4.0
SEQ ID NO 74298
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-74298

| | Query Match | 14.9% | Score 35.4 | DB 4 | Length 601 |
|----|-----------------------|--|-----------------|----------|------------|
| | Best Local Similarity | 63.5% | Pred. No. 0.036 | | |
| | Matches 54 | Conservative 0 | Mismatches 31 | Indels 0 | Gaps 0 |
| Qy | 105 | CTTCCCTTCATGCATCCATTGTTCTGTTTCATTCATTCATCATCCATCCATGCTCTG | 164 | | |
| Db | 158 | CATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCAACA | 217 | | |

Qy 165 AGCTTCACCTCTGACTCCCTAACT 189
| | | | | | | | | | | | | | | |
Db 218 AATTTTATTCTCTCTCTCTCT 242

Search completed: April 7, 2005, 15:00:38
Job time : 59.9231 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2005, 10:54:05 ; Search time 213.086 Seconds

(without alignments)
6768.525 Million cell updates/sec

Title: US-09-977-579-4_COPY_1024_1261

Perfect score: 238

Sequence: 1 aacaggagcagtgatgacatg.....tgccagaactgagaagccgg 238

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5615251 seqs, 3030001701 residues

Total number of hits satisfying chosen parameters: 11230502

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications NA:*
- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
 - 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
 - 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
 - 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
 - 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
 - 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
 - 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
 - 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
 - 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
 - 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
 - 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
 - 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
 - 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
 - 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
 - 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
 - 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
 - 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
 - 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
 - 19: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
 - 20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
 - 21: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
 - 22: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query Match | Score | Length | ID | Description |
|------------|-------------|-------|--------|----|----------------------|
| 1 | 238 | 100.0 | 1261 | 11 | US-09-977-579-4 |
| 2 | 216 | 90.8 | 4625 | 10 | US-09-764-891-7659 |
| 3 | 40.6 | 17.1 | 221000 | 17 | US-10-174-014-12 |
| 4 | 40.6 | 17.1 | 233380 | 13 | US-10-087-192-652 |
| 5 | 38.4 | 16.1 | 228 | 18 | US-10-674-124A-21325 |
| 6 | 38 | 16.0 | 27079 | 17 | US-10-034-650-55 |
| 7 | 38 | 16.0 | 38360 | 10 | US-09-999-121-14 |
| c 8 | 37.8 | 15.9 | 109 | 18 | US-10-674-124A-4659 |
| c 9 | 37.8 | 15.9 | 195 | 18 | US-10-674-124A-4658 |
| c 10 | 37.8 | 15.9 | 1091 | 13 | US-10-027-632-30772 |
| c 11 | 37.8 | 15.9 | 1091 | 17 | US-10-027-632-30772 |

| | | | | | | |
|------|------|------|--------|----|----------------------|--------------------|
| 12 | 37.8 | 15.9 | 1327 | 13 | US-10-027-632-215913 | Sequence 215913, |
| 13 | 37.8 | 15.9 | 1327 | 17 | US-10-027-632-215913 | Sequence 215913, |
| c 14 | 37.8 | 15.9 | 3768 | 9 | US-09-969-708-464 | Sequence 464, App |
| c 15 | 37.8 | 15.9 | 3768 | 9 | US-09-954-531-164 | Sequence 164, App |
| c 16 | 37.8 | 15.9 | 3768 | 9 | US-09-954-531-373 | Sequence 373, App |
| c 17 | 37.8 | 15.9 | 3768 | 10 | US-09-960-706-1041 | Sequence 1041, App |
| c 18 | 37.8 | 15.9 | 3768 | 10 | US-09-873-313-689 | Sequence 689, App |
| c 19 | 37.8 | 15.9 | 3768 | 19 | US-10-843-641A-1231 | Sequence 1231, App |
| c 20 | 37.8 | 15.9 | 3768 | 19 | US-10-843-641A-1440 | Sequence 1440, App |
| c 21 | 37.8 | 15.9 | 3768 | 19 | US-10-843-641A-7935 | Sequence 7935, App |
| c 22 | 37.8 | 15.9 | 8222 | 19 | US-10-486-319A-6 | Sequence 6, Appli |
| c 23 | 37.8 | 15.9 | 8222 | 19 | US-10-486-319A-42 | Sequence 42, Appl |
| c 24 | 37.8 | 15.9 | 8222 | 19 | US-10-486-319A-64 | Sequence 64, Appl |
| c 25 | 37.6 | 15.8 | 339 | 18 | US-10-674-124A-8325 | Sequence 8325, App |
| c 26 | 37.6 | 15.8 | 1781 | 13 | US-10-027-632-200093 | Sequence 200093, |
| c 27 | 37.6 | 15.8 | 1781 | 17 | US-10-027-632-200093 | Sequence 200093, |
| c 28 | 37.2 | 15.6 | 775062 | 18 | US-10-719-993-6844 | Sequence 6844, App |
| c 29 | 37 | 15.5 | 2632 | 13 | US-10-029-191-22 | Sequence 22, Appl |
| c 30 | 37 | 15.5 | 3108 | 13 | US-10-029-191-1 | Sequence 1, Appli |
| c 31 | 37 | 15.5 | 6419 | 15 | US-10-311-455-239 | Sequence 239, App |
| c 32 | 36.8 | 15.5 | 434 | 18 | US-10-674-124A-23074 | Sequence 23074, A |
| c 33 | 36.6 | 15.4 | 148 | 18 | US-10-674-124A-24954 | Sequence 24954, A |
| c 34 | 36.6 | 15.4 | 494 | 13 | US-10-027-632-320641 | Sequence 320641, |
| c 35 | 36.6 | 15.4 | 494 | 17 | US-10-027-632-320641 | Sequence 320641, |
| c 36 | 36.6 | 15.4 | 642 | 13 | US-10-027-632-61529 | Sequence 61529, A |
| c 37 | 36.6 | 15.4 | 642 | 17 | US-10-027-632-61529 | Sequence 61529, A |
| c 38 | 36.6 | 15.4 | 4743 | 10 | US-09-764-891-7898 | Sequence 7898, App |
| c 39 | 36.6 | 15.4 | 4747 | 10 | US-09-764-891-7897 | Sequence 7897, App |
| c 40 | 36.4 | 15.3 | 29322 | 17 | US-10-034-650-58 | Sequence 58, Appl |
| c 41 | 36.4 | 15.3 | 140152 | 18 | US-10-684-422-66 | Sequence 66, Appl |
| c 42 | 36.2 | 15.2 | 615 | 13 | US-10-027-632-278139 | Sequence 278139, |
| c 43 | 36.2 | 15.2 | 615 | 17 | US-10-027-632-278139 | Sequence 278139, |
| c 44 | 36.2 | 15.2 | 626 | 13 | US-10-027-632-278138 | Sequence 278138, |
| c 45 | 36.2 | 15.2 | 626 | 17 | US-10-027-632-278138 | Sequence 278138, |

ALIGNMENTS

RESULT 1

- US-09-977-579-4
- Sequence 4, Application US/09977579
- Publication No. US20040248240A1
- GENERAL INFORMATION:
- APPLICANT: Cambridge University Technical Services
- TITLE OF INVENTION: A novel family of beta sub-unit proteins from a voltage gated sod
- TITLE OF INVENTION: channel
- FILE OF INVENTION: nucleic acids encoding them and therapeutic or diagnostic uses ti
- FILE REFERENCE: 674558-2001
- CURRENT APPLICATION NUMBER: US/09/977,579
- CURRENT FILING DATE: 2001-10-15
- PRIOR APPLICATION NUMBER: PCT/EP00/01783
- PRIOR FILING DATE: 2000-02-24
- PRIOR APPLICATION NUMBER: 60,129,473
- PRIOR FILING DATE: 2000-02-24
- NUMBER OF SEQ ID NOS: 47
- SOFTWARE: Patentin version 3.1
- SEQ ID NO 4
- LENGTH: 1261
- TYPE: DNA
- ORGANISM: Homo sapiens
- US-09-977-579-4

| | | | | |
|-----------------------|--------------|--|---------------|-------------------|
| Query Match | 100.0% | Score 238; | DB 11; | Length 1261; |
| Best Local Similarity | 100.0%; | Pred. No. 9.2e-69; | | |
| Matches 238; | Conservative | 0; | Mismatches 0; | Indels 0; Gaps 0; |
| QY | 1 | AACAGGAGCAGTGTGACATGAGTGCCTGAACACCTGAGGAGCTGGACATCCCATCTTC | 60 | |
| Db | 1024 | AACAGGAGCAGTGTGACATGAGTGCCTGAACACCTGAGGAGCTGGACATCCCATCTTC | 1083 | |
| QY | 61 | AGCAATGTCAATGCCATCAGGAGGCGCCCAAGAGGCCCATCGCTTCCCTTCATCATC | 120 | |

Db 1084 AGCAATGTCATGGCATCAGAGGGGCGCCCAAGGGCCCATCGCTTCCCTTCATGCATC 1143
QY 121 CATTGTTCTGTTCAATTCATTCATCATCATCCACCTGCCTCTGAGCTTTCACCTCTGAC 180
Db 1144 CATTGTTCTGTTCAATTCATTCATCATCATCCACCTGCCTCTGAGCTTTCACCTCTGAC 1203
QY 181 TCCCTAACTCCATCAGACCTCTACGCACCACTAAGACTCTGCCAGAACTGAGAAGCGG 238
Db 1204 TCCCTAACTCCATCAGACCTCTACGCACCACTAAGACTCTGCCAGAACTGAGAAGCGG 1261

RESULT 2

US-09-764-891-7659
; Sequence 7659, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7659
; LENGTH: 4625
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-7659

Query Match 90.8%; Score 216; DB 10; Length 4625;
Best Local Similarity 100.0%; Pred. No. 2.9e-61;
Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 21 AGGTGCGCTGAACACCTGAGGGAGCTGGACATCCCATGTTTCAGCAATGTCAATGGCATCAG 80
Db 33 AGGTGCGCTGAACACCTGAGGGAGCTGGACATCCCATGTTTCAGCAATGTCAATGGCATCAG 92
QY 81 GAGGGCGCCCCAAGGGCCCCATCGCTTCCCTTCATGCATCCATTTGTTCTGTTCAATTCATT 140
Db 93 GAGGGCGCCCCAAGGGCCCCATCGCTTCCCTTCATGCATCCATTTGTTCTGTTCAATTCATT 152
QY 141 CATCATATACATCCACCTGCCTCTGAGCTTTCACCTTCATGCATCCATCCCAATCCATCAGACCT 200
Db 153 CATCATATACATCCACCTGCCTCTGAGCTTTCACCTTCATGCATCCATCCCAATCCATCAGACCT 212
QY 201 CTACGCACCATTAAGACTCTGCCAGAACTGAGAAGCC 236
Db 213 CTACGCACCATTAAGACTCTGCCAGAACTGAGAAGCC 248

RESULT 3

US-10-174-014-12
; Sequence 12, Application US/10174014
; Publication No. US20040005292A1
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Susan M. Freier
; APPLICANT: Kenneth W. Doble
; TITLE OF INVENTION: ANTISENSE MODULATION OF SMRT EXPRESSION
; FILE REFERENCE: PTS-0012
; CURRENT APPLICATION NUMBER: US/10/174,014
; CURRENT FILING DATE: 2002-06-17
; NUMBER OF SEQ ID NOS: 73
; SEQ ID NO 12
; LENGTH: 221000
; TYPE: DNA
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77967
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure

; NAME/KEY: unsure
; LOCATION: 77968
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77969
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77970
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77971
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77972
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77973
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77974
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77975
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77976
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77977
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77978
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77979
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77980
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77981
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77982
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77983
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77984
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77985
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure


```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(233380)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-652

Query Match      17.1%; Score 40.6; DB 13; Length 233380;
Best Local Similarity 57.5%; Pred. No. 0.027;
Matches 73; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 46 GGACATCCCATGTTAGCAATGTCATGCGATCAGAGGCGCCCAAGGGCCCATCGC 105
DB 55725 GGAGTCCCATGCTGGGAAGATTCATTCATTTGGGAGTCCCATCTGGGAGATCC 55784

QY 106 TTCCCTTCATGCATCCATGTTCTGTGTTTCATTCATCCATACATCCACCTGCTCTGA 165
DB 55785 ATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCTGGCAATA 55844

QY 166 GCTTCA 172
DB 55845 CTTATCA 55851

RESULT 5
US-10-674-124A-21325
; Sequence 21325, Application US/10674124A
; Publication No. US20040197797A1
; GENERAL INFORMATION:
; APPLICANT: INOKO, Hidetoshi
; APPLICANT: TAMIYA, Gen
; TITLE OF INVENTION: GENE MAPPING METHOD USING MICROSATELLITE
; FILE REFERENCE: GENETIC POLYMORPHISM MARKERS
; CURRENT APPLICATION NUMBER: US/10/674,124A
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 10/257,511
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: PCT/JP00/07621
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: JP2000-112699
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: JP2002-327516
; PRIOR FILING DATE: 2002-09-28
; PRIOR APPLICATION NUMBER: JP2002-383869
; PRIOR FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 27110
; SEQ ID NO 21325
; LENGTH: 228
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: AC009700.4_30060
; FEATURE:
; OTHER INFORMATION: Located on chromosome 15
; FEATURE:
; OTHER INFORMATION: Distance between a terminus base of telomere on
; OTHER INFORMATION: chromosomal short arm and 5'-terminus of this base
; OTHER INFORMATION: sequence : 41746864
; FEATURE:
; OTHER INFORMATION: Distance between 3'-terminus of neighbour sequence of
; OTHER INFORMATION: sequence listing upward to telomere on chromosomal short arm and
; OTHER INFORMATION: 5'-terminus of this base sequence : 51760
US-10-674-124A-21325

Query Match      16.1%; Score 38.4; DB 18; Length 228;
Best Local Similarity 62.5%; Pred. No. 0.018;
Matches 60; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 105 CTTCCCTTCATGCATCCATGTTCTGTTCATTCATCCATCCATCCATCCATCCGCTCG 164
DB 111 CATCCATCCATCCATCCATCCATCCATTCATTCATTCATTCATTCATTCATTCATTC 170
```

```
QY 165 AGCTTTACCTCTGACTCCCTAACTCCATCCACCT 200
DB 171 AACTAGCATCTCTTAGTTCTTAAGTCATTGGGGCT 206

RESULT 6
US-10-034-650-55
; Sequence 55, Application US/10034650
; Publication No. US20030216558A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David
; APPLICANT: Engelhard, Eric
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000128
; CURRENT APPLICATION NUMBER: US/10/034,650
; CURRENT FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: US 09/474,377
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 55
; LENGTH: 27079
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-034-650-55

Query Match      16.0%; Score 38; DB 17; Length 27079;
Best Local Similarity 57.6%; Pred. No. 0.1;
Matches 68; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 104 GCTTCCTTCATGCATCCATGTTCTGTTCATTCATCCATCCATCCACCTGCTCT 163
DB 3703 GCTTCCTTCATGCATCCATGTTCTGTTCATTCATCCATCCATCCATCCATCCAT 3762

QY 164 GAGCTTTACCTCTGACTCCCTAACTCCATCCATCCATCCATCCATCCATCCATCCAT 221
DB 3763 CCAACATCCATCTGTCTCCCTCCACCCACCATCATCTACACATCTCTCCCTCCATG 3820

RESULT 7
US-09-999-121-14
; Sequence 14, Application US/09999121
; Publication No. US20030039982A1
; GENERAL INFORMATION:
; APPLICANT: Ryan, James W.
; TITLE OF INVENTION: ISOLATED GENOMIC NUCLEOTIDE FRAGMENTS FROM THE p15 REGION OF
; FILE REFERENCE: CHROMOSOME 11
; CURRENT APPLICATION NUMBER: US/09/999,121
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: US 60/244,705
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 14
; LENGTH: 38360
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-999-121-14

Query Match      16.0%; Score 38; DB 10; Length 38360;
Best Local Similarity 62.8%; Pred. No. 0.11;
Matches 59; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 97 CCCATCCCTTCCTTCATGCATCCATGTTCTGTTCATTCATCCATCCATCCATCCATCC 156
DB 25309 CCCCTTTCTTCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCCATCC 25368

QY 157 TGCCTCTGAGCTTTTCACCTCTGACTCCCTTAATC 190
```


| | | | | | | | |
|-----------------------|-------|--------------|-------|------------|----|--------|------|
| Query Match | 15.9% | Score | 37.8 | DB | 9 | Length | 3768 |
| Best Local Similarity | 60.0% | Pred. No. | 0.066 | | | | |
| Matches | 63 | Conservative | 0 | Mismatches | 42 | Indels | 0 |
| | | | | | | Gaps | 0 |

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2005, 08:50:34 ; Search time 1419.88 Seconds
(without alignments)
6380.306 Million cell updates/sec

Title: US-09-977-579-4_COPY_1024_1261

Perfect score: 238

Sequence: 1 aacaggagcagtgtgacatg.....tgccagaactgagaagccgg 238

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hc:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gss1:*

9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 236 | 99.2 | 975 | 5 | BX452112 |
| 2 | 236 | 99.2 | 2555 | 3 | CR609664 |
| 3 | 236 | 99.2 | 4052 | 3 | HS801563 |
| 4 | 214.4 | 90.1 | 413 | 7 | D4825 |
| 5 | 158.6 | 66.6 | 926 | 5 | BX411288 |
| 6 | 101.2 | 42.5 | 484 | 6 | CB441198 |
| 7 | 95.6 | 40.2 | 1062 | 5 | BX420015 |
| 8 | 46.4 | 19.5 | 637 | 3 | CE517323 |
| 9 | 42.2 | 17.7 | 4149 | 3 | AK049286 |
| 10 | 42 | 17.6 | 322 | 4 | EG001495 |
| 11 | 41.8 | 17.6 | 601 | 6 | CB442219 |
| 12 | 40.8 | 17.1 | 941 | 9 | CNS02DGR |
| 13 | 40.2 | 16.9 | 844 | 9 | CNS02UEH |
| 14 | 40 | 16.8 | 611 | 8 | BH308136 |
| 15 | 39.4 | 16.6 | 702 | 9 | CR196841 |
| 16 | 39.2 | 16.5 | 676 | 8 | AQ835006 |
| 17 | 39 | 16.4 | 564 | 8 | AZ627327 |
| 18 | 39 | 16.4 | 697 | 9 | CE47269 |
| 19 | 38.8 | 16.3 | 712 | 5 | BP380503 |
| 20 | 38.6 | 16.2 | 660 | 6 | BY730622 |
| 21 | 38.6 | 16.2 | 675 | 9 | CE070920 |
| 22 | 38.6 | 16.2 | 1101 | 9 | CNS0512Z |
| 23 | 38.4 | 16.1 | 668 | 9 | AG087943 |
| 24 | 38.4 | 16.1 | 750 | 8 | BH356347 |

| | | | | | |
|----|------|------|------|---|----------|
| 25 | 38.2 | 16.1 | 694 | 9 | CE244764 |
| 26 | 38.2 | 16.1 | 733 | 9 | AG432316 |
| 27 | 38.2 | 16.1 | 751 | 8 | AZ957360 |
| 28 | 38 | 16.0 | 835 | 8 | AQ748173 |
| 29 | 38 | 16.0 | 1138 | 1 | AL564828 |
| 30 | 37.8 | 15.9 | 537 | 7 | CF979338 |
| 31 | 37.8 | 15.9 | 982 | 9 | CNS05CTG |
| 32 | 37.8 | 15.9 | 1045 | 9 | CNS05M4G |
| 33 | 37.6 | 15.8 | 221 | 9 | CE443630 |
| 34 | 37.6 | 15.8 | 436 | 8 | AQ813992 |
| 35 | 37.6 | 15.8 | 445 | 8 | AQ201281 |
| 36 | 37.4 | 15.7 | 372 | 2 | BE181389 |
| 37 | 37.4 | 15.7 | 542 | 8 | BZ173993 |
| 38 | 37.4 | 15.7 | 543 | 7 | CV396629 |
| 39 | 37.2 | 15.6 | 323 | 9 | CE484391 |
| 40 | 37.2 | 15.6 | 443 | 7 | CN315915 |
| 41 | 37.2 | 15.6 | 497 | 5 | BX258128 |
| 42 | 37.2 | 15.6 | 530 | 8 | BZ689975 |
| 43 | 37.2 | 15.6 | 686 | 7 | CO377943 |
| 44 | 37.2 | 15.6 | 737 | 9 | CE805346 |
| 45 | 37.2 | 15.6 | 743 | 9 | CR222933 |

ALIGNMENTS

RESULT 1
BX452112
LOCUS BX452112 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
DEFINITION CS0DF023YA09 5-PRIME, mRNA sequence.
ACCESSION BX452112 GI:47054503
VERSION BX452112.2
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 975)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 22, 2003 this sequence version replaced gi:31024339.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized. Library was constructed by Life Technologies, a division of Invitrogen.
This sequence belongs to sequence cluster 6147.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?8=CS0BAG018ZF03_CS01654_l&c=6147.r

FEATURES

Location/Qualifiers
1..975
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF023YA09"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/note="Organ: Brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

Query Match 99.2%; Score 236; DB 5; Length 975;

Best Local Similarity 100.0%; Pred. No. 1.1e-59; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACAGGAGCAGTGTGACATGAGTGGCTGAAACCTGAGGGAGCTGGACATCCCATGTTTC 60

Db 117 AACAGGAGCAGTGTGACATGAGTGGCTGAAACCTGAGGGAGCTGGACATCCCATGTTTC 176

QY 61 AGCAATGTCAATGGCATCAGAGGGGCGCCCAAGAGGCCCATCGCTTCCCTTCATGCATC 120

Db 177 AGCAATGTCAATGGCATCAGAGGGGCGCCCAAGAGGCCCATCGCTTCCCTTCATGCATC 236

QY 121 CATTGTCTCTGTCATTCATCCATACATCACCCTGCTCTGAGCTTTCACCTCTGAC 180

Db 237 CATTGTCTCTGTCATTCATCCATACATCACCCTGCTCTGAGCTTTCACCTCTGAC 296

QY 181 TCCTTAACCTCCATCAGACCTCTACGACCATAGACTCTGCCAGAACTGAGAAGCC 236

Db 297 TCCTTAACCTCCATCAGACCTCTACGACCATAGACTCTGCCAGAACTGAGAAGCC 352

RESULT 2
LOCUS CR609664 2555 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CS0DF023YA09 of Fetal brain of Homo sapiens (human).
ACCESSION CR609664
VERSION CR609664.1 GI:50490471
KEYWORDS HTC; CDS; CDNA
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished

Contact: Feng Liang Email: fliang@lifetech.com URL: http://fulllength.invitrogen.com/

REFERENCE 2 (bases 1 to 2555)
Genoscope.

Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail: seqref@genoscope.cns.fr)

- Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library

was normalized. Library was constructed by Life Technologies, a

division of Invitrogen.

FEATURES
Location/Qualifiers
1..2555

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DF023YA09"

/tissue_type="Fetal brain"

/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 99.2%; Score 236; DB 3; Length 2555;

Best Local Similarity 100.0%; Pred. No. 1.4e-59; Mismatches 0; Indels 0; Gaps 0;

Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACAGGAGCAGTGTGACATGAGTGGCTGAAACCTGAGGGAGCTGGACATCCCATGTTTC 60

Db 883 AACAGGAGCAGTGTGACATGAGTGGCTGAAACCTGAGGGAGCTGGACATCCCATGTTTC 942

QY 61 AGCAATGTCAATGGCATCAGAGGGGCGCCCAAGAGGCCCATCGCTTCCCTTCATGCATC 120

Db 943 AGCAATGTCAATGGCATCAGAGGGGCGCCCAAGAGGCCCATCGCTTCCCTTCATGCATC 1002

QY 121 CATTGTCTCTGTCATTCATCCATACATCACCCTGCTCTGAGCTTTCACCTCTGAC 180

Db 1003 CATTGTCTCTGTCATTCATCCATACATCACCCTGCTCTGAGCTTTCACCTCTGAC 1062

QY 181 TCCTTAACCTCCATCAGACCTCTACGACCATAGACTCTGCCAGAACTGAGAAGCC 236

Db 1063 TCCTTAACCTCCATCAGACCTCTACGACCATAGACTCTGCCAGAACTGAGAAGCC 1118

RESULT 3
LOCUS HSM801563 4052 bp mRNA linear HTC 23-SEP-2004
DEFINITION Homo sapiens mRNA; cDNA DKP2p761F182 (from clone DKP2p761F182).
ACCESSION AL136589
VERSION AL136589.1 GI:13276680
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Ottenwaelder, B., Obermaier, B., Deutschenbaur, S., Schaipp, A.,

Mewes, H.W., Weil, B., Amid, C., Osanger, A., Fobo, G., Han, M. and

Wiemann, S.

The German cDNA Consortium

Direct Submission

Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764

Neuherberg, GERMANY

Research Center (DKPZ); Email s.wiemann@dkfz-heidelberg.de;

sequenced by Medigenomix (Martinsried/Germany) within the cDNA

sequencing consortium of the German Genome Project.

This clone (DKP2p761F182) is available at the RZPD Deutsches

Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.

Please contact RZPD for ordering:

http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKP2p761F182

Further information about the clone and the sequencing project is

available at http://mips.gsf.de/projects/cdna/.

FEATURES
Location/Qualifiers
1..4052

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="RZPD:DKP2p761F182"

/db_xref="taxon:9606"

/clone="DKP2p761F182"

/tissue_type="amygdala"

/clone_lib="761 (synonym: hamy2). Vector pSport1; host

DH10B; sites NotI + SalI"

/dev_stage="adult"

/note="voltage-gated sodium channel beta-3 subunit"

1..4052

/gene="DKP2p761F182"

804..1451

/gene="DKP2p761F182"

/codon_start=1

/product="hypothetical protein"

/protein_id="CAB68524.1"

/db_xref="GI:13276681"

/db_xref="GOA:Q9NY72"

/translation="MPAFNRLPLASLVLIYVSVCFVVCVVPSETEAVQGNPMKLR

CISCMKREVEATTVVEFYRPEGGKDFLIYVRNGHVEVSPFOGRLQWNGSKDLQD

VSITVNTLDSGLYTCNVSEFEFAHRPFVKTRLIPLRVTEAGDFTSVVSSEI

MMYLLVFLTLMLLEMIYCYRKVSKAEAEANASDYLAIIFSENKENSAPVVEE"

ORIGIN
Query Match 99.2%; Score 236; DB 3; Length 4052;

Best Local Similarity 100.0%; Pred. No. 1.5e-59; Mismatches 0; Indels 0; Gaps 0;

Matches 236; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACAGGAGCAGTGTGACATGAGTGGCTTGAACACCTGAGGGAGCTGGACATCCCATGTTTC 60

Db 1452 AACAGGAGCAGTGTGACATGAGTGGCTTGAACACCTGAGGGAGCTGGACATCCCATGTTTC 1511

| | | | | | |
|------------|--|-------------|------|--------|-----------------|
| LOCUS | EX4111288 | 926 bp | mRNA | linear | EST 03-MAY-2004 |
| DEFINITION | EX4111288 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone CS0D7023YA09 3-PRIME, mRNA sequence. | | | | |
| ACCESSION | EX4111288 | | | | |
| VERSION | EX4111288.1 | GI:30765121 | | | |
| KEYWORDS | EST. | | | | |
| SOURCE | Homo sapiens (human) | | | | |
| ORGANISM | Homo sapiens | | | | |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | | |

full-length cDNA libraries and normalization
 Unpublished (2001)
 Contact: Genoscope
 Genoscope - Centre National de Séquençage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
 was not normalized. Library was constructed by Life Technologies, a
 division of Invitrogen.

```

FEATURES
source
    Location/Qualifiers
    1..926
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="CSODF023YA09"
    /tissue_type="FETAL BRAIN"
    /dev_stage="fetal"
    /clone_lib="Homo sapiens FETAL BRAIN"
    /notes="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
    was primed with a NotI-oligo(dT) primer. Five prime end
    enriched, double-strand cDNA was digested with Not I and
    cloned into the Not I and EcoRV sites of the pCMVSPORT 6

```

```

ORIGIN
Query Match      66.6%; Score 158.6; DB 5; Length 926;
Best Local Similarity 87.4%; Pred. No. 1.8e-36;
Matches 208; Conservative 0; Mismatches 25; Indels 5; Gaps 3;
ov 1 AACAGGACGAGTCTGACATGAGTGTGGCTTGAACACCTGAGGGACTGCACATCCCATGTTTC 60

```

[illegible]

| | | | | |
|------------|---|--------|--------|-----------------|
| LOCUS | CB441198 | 484 bp | linear | EST 25-MAR-2003 |
| DEFINITION | 691539 MARC 6BOV Bos taurus cDNA 5', mRNA sequence. | | | |
| ACCESSION | CB441198 | | | |

| Source | Organism |
|------------------|------------|
| Bos taurus (cow) | Bos taurus |

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.

REFERENCE
AUTHORS Smith,T.P.L., Roberts,A.J., Echternkamp,S.E., Chitko-McKown,C.G., Wray,J.E. and Keele,J.W.
TITLE A second set of bovine ESTs from pooled-tissue normalized libraries
JOURNAL Unpublished (2003)
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@mail.marc.usda.gov

Single pass sequencing. Bases called with phred v0.020425.c and trimmed with the aid of the trim_alt option. Vector identified with cross_match v0.990329.
Plate: FOY8048 row: K column: 21
Seq primer: GTAATACGACTCATTATAGG.

FEATURES
source
1. .484
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 6BOV"
/note="Vector: pcDNA3.1; Site 1: EcoRI; Site 2: NotI; Library made with RNA pooled from multiple tissues including liver, lung, hypothalamus, pituitary, and placenta/endometrium."

ORIGIN
Query Match 42.5%; Score 101.2; DB 6; Length 484;
Best Local Similarity 78.5%; Pred. No. 2.6e-19;
Matches 135; Conservative 0; Mismatches 33; Indels 4; Gaps 1;

QY 65 ATGTCATGCGATCAGGAGGGGCCCAAGGGCCCATCGCTTCCCTTCATGATTCATT 124
DB 4 ATATCAACTGCGATCAGGAGGGTCCCTAGGACCATCACTTCCCTTCGTCGTCATC 63

QY 125 GTTCTGTTTCATTCATTCATCATACATCACTGCTGCTGAGCTTTCACCTCTGACTCC 184
DB CT-----TTGAGTTTCATTCATTCATCATCACTGCTGAGCTTTCACCTCTGACTCC 119

QY 185 TTAATCCATCAGACCTCTACGACCACTAAGACTTCCGCAAGACTGAGAGGCC 236
DB 120 TTAATCCACAGACCTCTACACACCTCAAGACTTTGCCAGAACTGAGGAGCC 171

RESULT 7
BX420015 1062 bp mRNA linear EST 01-MAY-2004
LOCUS BX420015 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
DEFINITION CS09F023YA09 5-PRIME, mRNA sequence.

ACCESSION BX420015
VERSION BX420015.2 GI:46929710
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 13, 2003 this sequence version replaced gi:30646738.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized. Library was constructed by Life Technologies, a division of Invitrogen.

This sequence belongs to sequence cluster 6147.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS09F023AA05QP1&c=6147.r.

FEATURES
Location/Qualifiers
1. .1062
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS09F023YA09"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN
Query Match 40.2%; Score 95.6; DB 5; Length 1062;
Best Local Similarity 68.1%; Pred. No. 1.5e-17;
Matches 124; Conservative 14; Mismatches 43; Indels 1; Gaps 1;

QY 1 AACAGAGCAGTGTGACATGAGTGGCGCTGAAACACTGAGGAGCTGGACATCCCATGTTTC 60
DB 882 AACAGAGCAGTGTGACATGAGTGGCGCTGAAACACTGAGGAGCTGGACATCCCATGTTTC 941

QY 61 ACATATGTCATGATCAGGAGGGCGCCCAAGGGCCCATCGCTTCCCTTCATGATC 120
DB 942 ACAAATGTAATGATGAGGAGGGCGCCCAAGGG-CCCATCSCCTTCTTTCATGATC 1000

QY 121 CATGTGTTCTGTTTCATTCATTCATCATCATCATCCATCGCTCTGAGCTTTCACCTCTGAC 180
DB 1001 ATTGTYTTTTCATTATTATTCACAAATCCTKCTTCTTACGTTTACGTTTCTTACAA 1060

QY 181 TC 182
DB 1061 TC 1062

RESULT 8
CES17323 637 bp DNA linear GSS 28-SEP-2003
LOCUS tigr-gss-dog-17000327420279 Dog Library Canis familiaris genomic,
DEFINITION genomic survey sequence.

ACCESSION CES17323
VERSION CES17323.1 GI:36834104
KEYWORDS GSS.
SOURCE Canis familiaris (dog)

ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

REFERENCE
AUTHORS Kirkness,E.F., Baina,V., Halpern,A.L., Levy,S., Remington,K., Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and Venter,J.C.
TITLE The dog genome: survey sequencing and comparative analysis
JOURNAL Science 301 (5641), 1898-1903 (2003)
MEDLINE 22875432
PUBMED 14512627

COMMENT The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirknes@tigr.org
Class: shotgun.
Location/Qualifiers
1. .637


```
/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/notes="Site 1: BstXI; Libraries were prepared from
peripheral blood"

ORIGIN

Query Match      19.5%; Score 46.4; DB 9; Length 637;
Best Local Similarity 59.8%; Pred. No. 0.0076;
Matches 80; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 33 CACCTGAGGACCTGGACATCCCATGTTTCAGCAATGTCATGCGATCAGAGGGCGCCCA 92
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 24 CAAGCGAAGCTGAGACACCCCAAGATGCTCCCAACTGGGTTCTGAAGGCTTCACA 83
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 93 AGGGCCCCATCGCTCCCTTCATGCAATCCATGTTCTGTTTCATTCATTCATTCATCATC 152
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 84 AGTGCCTTCGTTTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 143
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 153 CACCTGCTCTGAGCT 168
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 144 CACAAGCCCCATTCCT 159
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 9
LOCUS AK049286
DEFINITION Mus musculus ES cells cDNA, RIKEN full-length enriched library,
clone: C330019I03 product: VOLTAGE-GATED SODIUM CHANNEL BETA-3
SUBUNIT (SODIUM CHANNEL BETA 3 SUBUNIT) homolog [Rattus
norvegicus], full insert sequence.
ACCESSION AK049286
VERSION AK049286.1 GI:26093400
KEYWORDS HTc; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636

REFERENCE
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159

REFERENCE
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamanoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861

REFERENCE
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
PANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
MEDLINE 11076861

REFERENCE
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
PANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
MEDLINE 11076861
```

```
AUTHORS The PANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6 (bases 1 to 4149)
AUTHORS Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, P., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akai, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
TITLE Direct Submission
JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
FEATURES
source Location/Qualifiers
1..4149
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM.DB:C330019I03"
/db_xref="taxon:10090"
/clone="C330019I03"
/cell_type="ES cells"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
misc_feature 410..1058
/note="VOLTAGE-GATED SODIUM CHANNEL BETA-3 SUBUNIT (SODIUM
CHANNEL BETA 3 SUBUNIT) homolog [Rattus norvegicus]
(SPTR|Q9JUK0, evidence: FASTA, 99.5%ID, 100%length,
match=646)
putative"
4129..4134
/note="putative"
polyA_signal
polyA_site 4149
/note="putative"
ORIGIN
Query Match 17.7%; Score 42.2; DB 3; Length 4149;
Best Local Similarity 67.8%; Pred. No. 0.21;
Matches 59; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 36 CTGAGGAGCTGGACATCCCATGTTTCAGCAATGTCATGCGATCAGAGGGCGCCCAAG 95
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1310 CTGAGCAACTGGCTATCCCGAGTTCCAGCAATGCCCAATCAGGAAGTGCCTCCAGGT 1369
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 96 GCCCATCGCTTCCTTCATGCATCCA 122
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1370 GTCCCAACATCATCTTCTTATCA 1396
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 10
LOCUS BG001495/c
DEFINITION MR3-GN0229-141100-003-f07 GN0229 Homo sapiens cDNA, mRNA sequence.
```

```

ACCESSION   BG001495
VERSION     BG001495.1  GI:12439883
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 322)
AUTHORS     Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
            Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
            Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
            Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,
            O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
            Simpson, A.J.
TITLE       Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
JOURNAL     Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE    20202663
PUBMED     10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?cl=MR3&t2=MR3-GN0229-
            141100-003-f07&t3=2000-11-14&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 51
            High quality sequence stop: 161.
            Location/Qualifiers
            1..322
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /dev_stage="Adult"
            /clone_lib="GN0229"
            /note="Organ: placenta normal; Vector: puc18; Site 1:
            SmaI; Site 2: SmaI; A mini-library was made by cloning
            products derived from ORESTES PCR (U.S. Letters Patent
            application No. 196,716 - Ludwig Institute for Cancer
            Research) profiles into the pUC 18 vector. Reverse
            transcription of tissue mRNA and cDNA amplification were
            performed under low stringency conditions."
            ORIGIN
            Query Match 17.6%; Score 42; DB 4; Length 322;
            Best Local Similarity 64.3%; Pred. No. 0.14;
            Mismatches 63; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
            QY 98 CCCATCGCTTCCTTCATCCATCCATCTGTTGTTTCATTCATCCATCATCCACCT 157
            DB 118 CCTATCCCATTCATCCATCCATCCATCTACCTATCCATCCATCCATCCATCCACTT 59
            QY 158 GCCTCTGAGCTTTCACCTCTGACTCCCTAACTCCATCA 195
            DB 58 ACCTATCATTCATCCATCCATCCATCCATCATCTCCATCA 21
            RESULT 11
            LOCUS      CB442219/c
            DEFINITION 692787 MARC 6BOV Bos taurus cDNA 3', mRNA sequence.
            ACCESSION  CB442219
            VERSION     CB442219.1  GI:29230022
            SOURCE      Bos taurus (cow)
            ORGANISM    Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
            Bovinae; Bos.
            1 (bases 1 to 601)
            Smith, T.P.L., Roberts, A.J., Echternkamp, S.E., Chitko-McKown, C.G.,
            Wray, J.E. and Keele, J.W.
            A second set of bovine ESTs from pooled-tissue normalized libraries
            Unpublished (2003)
            Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called with phred v0.020425.c and
            trimmed with the aid of the trim_alt option. Vector identified with
            cross_match v0.990329.
            Plate: FOY8048 row: K column: 21
            Seq primer: TAGAAGGCACAGCTCGAGG.
            Location/Qualifiers
            1..601
            /organism="Bos taurus"
            /mol_type="mRNA"
            /db_xref="taxon:9913"
            /tissue_type="pooled"
            /lab_host="DH10B"
            /clone_lib="MARC 6BOV"
            /note="Vector: pCDNA3.1; Site 1: EcoRI; Site 2: NotI;
            Library made with RNA pooled from multiple tissues
            including liver, lung, hypothalamus, pituitary, and
            placenta/endometrium."
            ORIGIN
            Query Match 17.6%; Score 41.8; DB 6; Length 601;
            Best Local Similarity 86.8%; Pred. No. 0.18;
            Mismatches 46; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
            QY 184 CTAATCTCATCAGACCTCTACGCACATAGACTCTGCCAGACTCAGAGGCC 236
            DB 599 CTAATCTCACCAGACCTTTACACACTTTCACAGACTTTCACAGACTCAGAGGCC 547
            RESULT 12
            LOCUS      CNS02DGR/c
            DEFINITION Tetraodon nigroviridis genome survey sequence T7 end of clone
            260306 of library G from Tetraodon nigroviridis, genomic survey
            sequence.
            ACCESSION  AL192420.1  GI:7830524
            VERSION     AL192420
            GSS: genome survey sequence.
            SOURCE      Tetraodon nigroviridis
            ORGANISM    Tetraodon nigroviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetraodontidae; Tetraodontidae; Tetraodon.
            1
            Roest Crolius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C.,
            Bernot, A., Fizames, C., Wincker, P., Brottier, P., Quetier, F.,
            Saurin, W. and Weissenbach, J.
            Estimate of human gene number provided by genome-wide analysis
            using Tetraodon nigroviridis DNA sequence
            Nat. Genet. 25 (2), 235-238 (2000)
            20296633
            10835645
            Roest Crolius, H., Jaillon, O., Dasilva, C., Ozouf-Costaz, C.,
            Fizames, C., Fischer, C., Bouneau, L., Billault, A., Quetier, F.,
            Saurin, W., Bernot, A. and Weissenbach, J.
            Characterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetraodon nigroviridis
            Genome Res. 10 (7), 939-949 (2000)
            20359837

```

PUBLISHED 10899143
REFERENCE 3 (bases 1 to 941)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (12-APR-2000) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr
COMMENT This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/tetraodon.
FEATURES
source
1. .941
Location/Qualifiers
/organism="Tetraodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="260G06"
/clone_lib="G"
/note="Genoscope sequence ID : COAG260BD03LP1-end : T7"
ORIGIN
Query Match 17.1%; Score 40.8; DB 9; Length 941;
Best Local Similarity 51.4%; Pred. No. 0.39;
Matches 108; Conservative 5; Mismatches 92; Indels 5; Gaps 1;
QY 4 AGGAGCAGTGTGACATGAGTGGCTGAACACCTGAGGGACTGACATCCCATGTTCCAGC 63
Db 868 AGTGGCGCGGACACAGTGTACAGACACCTGTGGAGGAGAACACAGAGTGGGA 809
QY 64 AATGTCAATGGCATFCAGGAGGGCGGCCCAAGGGCCCATCGCTTCCCTTCATGCATCAT 123
Db 808 CAGCAGGASGTGAGRGGGGGGCGACCGTCTCCATCCCTCCCTCCATCCATC--- 752
QY 124 TGTTCGTTTCATTCATTCATCCATCATCCACCTGCTCTGAGCTTCACCTGATCTCC 183
Db 751 --ATCCCTCCATCCATCCCTCCATCCATCCCTCCATCCATCCATCCCTCCCTCC 694
QY 184 CTAACCTCCATCAGACCTCTAGCACCATAA 213
Db 693 ATCCCTCCATCCATCCCTCCATCCATCCATCCATCCATCCATCCATCCATCCCTCC 664
RESULT 13
CNS02UEH/c 844 bp DNA linear GSS 01-SEP-2000
LOCUS Tetraodon nigroviridis genome survey sequence T7 end of clone
DEFINITION 167G23 of library G from Tetraodon nigroviridis, genomic survey sequence.
ACCESSION AL214370
VERSION AL214370.1 GI:7873189
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontoidea; Tetraodontidae; Tetraodon.
REFERENCE 1
AUTHORS Roest Crolius.H., Jaillon.O., Dasilva.C., Bouneau.L., Fisher.C., Bernot.A., Fizames.C., Wincker.P., Brottier.P., Quetier.F., Saurin.W. and Weissenbach.J.
TITLE Estimate of human gene number provided by genome-wide analysis using Tetraodon nigroviridis DNA sequence
JOURNAL Nat. Genet. 25 (2), 235-238 (2000)
MEDLINE 20296633
PUBLISHED 10835645
REFERENCE 2
AUTHORS Roest Crolius.H., Jaillon.O., Dasilva.C., Ozouf-Costaz.C., Fizames.C., Fischer.C., Bouneau.L., Billault.A., Quetier.F., Saurin.W., Bernot.A. and Weissenbach.J.
TITLE Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis
JOURNAL Genome Res. 10 (7), 939-949 (2000)

PUBLISHED 20359837
REFERENCE 10899143
AUTHORS 3 (bases 1 to 844)
TITLE Direct Submission
JOURNAL Submitted (12-APR-2000) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr
COMMENT This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/tetraodon.
FEATURES
source
1. .844
Location/Qualifiers
/organism="Tetraodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="167G23"
/clone_lib="G"
/note="Genoscope sequence ID : COAG167AD12LP1-end : T7"
ORIGIN
Query Match 16.9%; Score 40.2; DB 9; Length 844;
Best Local Similarity 56.1%; Pred. No. 0.58;
Matches 60; Conservative 7; Mismatches 40; Indels 0; Gaps 0;
QY 105 CTTCCCTTCATGCATCCATTTGTTCTGTTTCATTCATCCATCCATCCATCCCTCTG 164
Db 661 CATCTTCCATCCATCCATCCATTCATTCATTCATTCATTCATTCATTCATTCAT 602
QY 165 AGCTTTCACCTTCAGTCCCTTAACCTCCATCCATCCATCCATCCATCCATCCAT 211
Db 601 ATCTTTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATCC 555
RESULT 14
BH308136 611 bp DNA linear GSS 03-DEC-2001
LOCUS CH230-193A19, TJ CHORI-230 Segment 1 Rattus norvegicus genomic clone
DEFINITION CH230-193A19, genomic survey sequence.
ACCESSION BH308136
VERSION BH308136.1 GI:17233605
KEYWORDS GSS.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 611)
AUTHORS Zhao.S., Shetty.J., Shatsman.S., Teagave.G., Geer.K., Shivartebeyn.A., Gebregeorgis.E., Overton.L., Russell.D., Chen.D., Riggs.F., de Jong.P. and Fraser.C.M.
TITLE Rat BAC End Sequences from Library CHORI-230 EcoRI segment
JOURNAL Unpublished (1999)
COMMENT Other GSSs: CH230-193A19.TV
Contact: Shaving Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230
(http://www.chori.org/bacpac/rat230.htm). For BAC library availability, please contact Pieter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/or ering information.htm). BAC end plate: 193 row: A column: 19
Seq primer: SP6
Class: BAC ends.
Location/Qualifiers
source
1. .611

Search completed: April 7, 2005, 14:55:47
Job time : 1426.88 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2005, 07:44:40 ; Search time 1788.76 Seconds
(without alignments)
10158.276 Million cell updates/sec

Title: US-09-977-579-4_COPY_1_375

Perfect score: 375

Sequence: 1 cctcctctccgagctgagc.....gaagccgagcccccagaag 375

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------------|--------------------|
| 1 | 375 | 100.0 | 1261 | 6 AR359850 | AR359850 Sequence |
| 2 | 375 | 100.0 | 1261 | 6 AX039100 | AX039100 Sequence |
| 3 | 375 | 100.0 | 1261 | 9 HSA243396 | AJ243396 Homo sapi |
| 4 | 373 | 99.5 | 127347 | 2 AP000682 | AP000682 Homo sapi |
| 5 | 373 | 99.5 | 144833 | 2 AC063921 | AC063921 Homo sapi |
| 6 | 373 | 99.5 | 149800 | 2 AC021981 | AC021981 Homo sapi |
| 7 | 373 | 99.5 | 172546 | 2 AC024604 | AC024604 Homo sapi |
| 8 | 373 | 99.5 | 178169 | 9 AP002765 | AP002765 Homo sapi |
| 9 | 373 | 99.5 | 181471 | 9 AC069539 | AC069539 Homo sapi |
| 10 | 350 | 93.3 | 178431 | 2 AP002749 | AP002749 Homo sapi |
| 11 | 164 | 43.7 | 174285 | 2 AC145484 | AC145484 Lemur cat |
| 12 | 122 | 32.5 | 66980 | 2 AC013796 | AC013796 Homo sapi |
| 13 | 74 | 19.7 | 66980 | 2 AC013796 | AC013796 Homo sapi |
| 14 | 53 | 14.1 | 521 | 11 PM12B8G | AL684326 Penicilli |
| 15 | 51 | 13.6 | 134226 | 2 AC148923 | AC148923 Alligator |
| 16 | 46.4 | 12.4 | 148643 | 2 AC011871 | AC011871 Homo sapi |
| 17 | 45.6 | 12.2 | 216449 | 2 AC051616 | AC051616 Mus muscu |
| 18 | 45.4 | 12.1 | 140307 | 2 AC135353 | AC135353 Mus muscu |
| 19 | 45.4 | 12.1 | 238861 | 2 AC148331 | AC148331 Mus muscu |

| | | | | | | | |
|---|----|------|------|--------|----|-------------|--------------------|
| C | 20 | 45.4 | 12.1 | 239130 | 2 | AC079420 | Mus muscu |
| C | 21 | 44.8 | 11.9 | 11608 | 9 | HSINT2 | X14445 Human int-2 |
| C | 22 | 44.8 | 11.9 | 104202 | 9 | AP006345 | AP006345 Homo sapi |
| C | 23 | 44.6 | 11.9 | 58293 | 8 | AC096781 | AC096781 Oryza sat |
| C | 24 | 44.6 | 11.9 | 117157 | 8 | AC034258 | AC034258 Oryza sat |
| C | 25 | 44.6 | 11.9 | 306161 | 8 | AB017102 | AB017102 Oryza sat |
| C | 26 | 44.4 | 11.8 | 21481 | 2 | CR847893 | CR847893 Danio rer |
| C | 27 | 44.4 | 11.8 | 59364 | 2 | AC102347 | AC102347 Mus muscu |
| C | 28 | 44.2 | 11.8 | 3252 | 9 | HSTAF113 | Y11354 H.sapiens m |
| C | 29 | 44.2 | 11.8 | 110000 | 2 | BX255276_08 | Continuation (9 of |
| C | 30 | 44.2 | 11.8 | 146691 | 2 | BX927196 | BX927196 Danio rer |
| C | 31 | 44 | 11.7 | 85434 | 2 | AC066610 | AC066610 Homo sapi |
| C | 32 | 44 | 11.7 | 223971 | 2 | AC115753 | AC115753 Mus muscu |
| C | 33 | 43.8 | 11.7 | 152347 | 2 | AC021096 | AC021096 Homo sapi |
| C | 34 | 43.8 | 11.7 | 196950 | 2 | CR589874 | CR589874 Danio rer |
| C | 35 | 43.6 | 11.6 | 972 | 11 | PM12A12G | AL684288 Penicilli |
| C | 36 | 43.6 | 11.6 | 1094 | 11 | PM7G11B | AL685196 Penicilli |
| C | 37 | 43.6 | 11.6 | 53897 | 2 | AC101434 | AC101434 Mus muscu |
| C | 38 | 43.6 | 11.6 | 191983 | 9 | AF274858 | AF274858 Homo sapi |
| C | 39 | 43.4 | 11.6 | 33676 | 1 | SCARD1GN | X84374 Saccharothr |
| C | 40 | 43.4 | 11.6 | 150399 | 2 | AC013385 | AC013385 Homo sapi |
| C | 41 | 43.4 | 11.6 | 240411 | 2 | AC105428 | AC105428 Mus muscu |
| C | 42 | 43.2 | 11.5 | 61967 | 2 | AC100100 | AC100100 Mus muscu |
| C | 43 | 43.2 | 11.5 | 64970 | 2 | AC099958 | AC099958 Mus muscu |
| C | 44 | 43.2 | 11.5 | 208238 | 2 | CR385024 | CR385024 Danio rer |
| C | 45 | 43.2 | 11.5 | 239130 | 2 | AC079420 | AC079420 Mus muscu |

ALIGNMENTS

| | | | | | | | |
|------------|---|-------------|------------------------------------|---------|-----|--------|-----------------|
| RESULT 1 | AR359850 | AR359850 | Sequence 4 from patent US 6593565. | 1261 bp | DNA | linear | PAT 17-AUG-2003 |
| LOCUS | AR359850 | AR359850 | Sequence 4 from patent US 6593565. | 1261 bp | DNA | linear | PAT 17-AUG-2003 |
| DEFINITION | AR359850 | AR359850 | Sequence 4 from patent US 6593565. | 1261 bp | DNA | linear | PAT 17-AUG-2003 |
| ACCESSION | AR359850 | AR359850 | Sequence 4 from patent US 6593565. | 1261 bp | DNA | linear | PAT 17-AUG-2003 |
| VERSION | AR359850.1 | GI:33766660 | Sequence 4 from patent US 6593565. | 1261 bp | DNA | linear | PAT 17-AUG-2003 |
| KEYWORDS | Unknown. | | | | | | |
| SOURCE | Unknown. | | | | | | |
| ORGANISM | Unknown. | | | | | | |
| REFERENCE | 1 (bases 1 to 1261) | | | | | | |
| AUTHORS | Heslin, P. and Lynam, N.R. | | | | | | |
| TITLE | Vehicle interior rearview mirror assembly including an accessory-containing housing | | | | | | |
| JOURNAL | Patent: US 6593565-A 4 15-JUL-2003; | | | | | | |
| FEATURES | Location/Qualifiers | | | | | | |
| source | 1..1261 | | | | | | |
| organism | "/organism="unknown" | | | | | | |
| mol_type | "/mol_type="genomic DNA" | | | | | | |

| | | | | | | | |
|-----------------------|--------|--|----------|------------|----|--------|-------|
| Query Match | 100.0% | Score | 375; | DB | 6; | Length | 1261; |
| Best Local Similarity | 100.0% | Pred. No. | 1.4e-70; | | | | |
| Matches | 375; | Conservative | 0; | Mismatches | 0; | Indels | 0; |
| Gaps | 0; | | | | | | |
| Qy | 1 | CCCTCCCTCCGAGCTGAGCTTACCTCGGCGCAACGAGCGAGCGCGCGAGTGG | 60 | | | | |
| Db | 1 | CCCTCCCTCCGAGCTGAGCTTACCTCGGCGCAACGAGCGAGCGCGAGTGG | 60 | | | | |
| Qy | 61 | AAGCTGGAGTTCGGGGTGGGGGAGCGACTGTCCTGGTGTCTGAGCGCGCGGAG | 120 | | | | |
| Db | 61 | AAGCTGGAGTTCGGGGTGGGGGAGCGACTGTCCTGGTGTCTGAGCGCGCGGAG | 120 | | | | |
| Qy | 121 | GGCGGCGCGAGCGGTGATTCGGCTCCCTCGAATCGGAGAGTCCAGTGGGGTGG | 180 | | | | |
| Db | 121 | GGCGGCGCGAGCGGTGATTCGGCTCCCTCGAATCGGAGAGTCCAGTGGGGTGG | 180 | | | | |
| Qy | 181 | GGCCCCAAAGCCCCACCCCGCTCCAAAGCTCCAGGGCTCCCGAGGACCCGCTCG | 240 | | | | |
| Db | 181 | GGCCCCAAAGCCCCACCCCGCTCCAAAGCTCCAGGGCTCCCGAGGACCCGCTCG | 240 | | | | |
| Qy | 241 | GCCTTCTTCCGTCAGAAAGTCGCCCCCTGGGGGCGAGTTCGTCCTCCAAAGGGT | 300 | | | | |

```
Db 241 GGCCTTCCTTCGGTCAGAAAGTCGCCCTCGGGGCGAGTTCTGCCAAAGGTTTCCTCG 300
QY 301 AAAGAATCTGAGAGGGCGCAGTCTTGTAGCCGAGGAATCTCTGTGTAGCCCTTGAAGC 360
Db 301 AAAGAATCTGAGAGGGCGCAGTCTTGTAGCCGAGGAATCTCTGTGTAGCCCTTGAAGC 360
QY 361 CGCCAGCCCGCAGAAG 375
Db 361 CGCCAGCCCGCAGAAG 375

RESULT 2
AX039100
LOCUS Homo sapiens 1261 bp DNA linear PAT 18-NOV-2000
DEFINITION Sequence 4 from Patent WO063367.
ACCESSION AX039100
VERSION AX039100.1 GI:11229276
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Morgan, K., Jackson, A., and Morgan, K.
TITLE A novel family of beta sub-unit proteins from a voltage-gated sodium channel, nucleic acids encoding them and therapeutic or diagnostic uses there of
JOURNAL PATENT: WO 063367-A 4 26-OCT-2000; WARNER-LAMBERT COMPANY (US); Cambridge University Technical Services Limited (GB)
FEATURES
    source
    1..1261
    /organism="Homo sapiens"
    /mol_type="unassigned DNA"
    /db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 375; DB 6; Length 1261;
Best Local Similarity 100.0%; Pred. No. 1.4e-70;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCCTCCCTCCGAGTGTACCTTGGCGGAGCGGAGTGTCCGTGTGTGTGAGCGCGCGAGA 60
Db 1 CCCTCCCTCCGAGTGTACCTTGGCGGAGCGGAGTGTCCGTGTGTGTGAGCGCGCGAGA 60
QY 61 AAGCTGGAGTTCGGGGTGGCGGGAGCGGAGTGTCCGTGTGTGTGAGCGCGCGAGA 120
Db 61 AAGCTGGAGTTCGGGGTGGCGGGAGCGGAGTGTCCGTGTGTGTGAGCGCGCGAGA 120
QY 121 GCGGCGCGGAGCGGTGATCGGCTCCCTCGAACTGGGGAGGTCCAGTGGGGTCCGCTTAG 180
Db 121 GCGGCGCGGAGCGGTGATCGGCTCCCTCGAACTGGGGAGGTCCAGTGGGGTCCGCTTAG 180
QY 181 GGCCTCAAGCCCGACCGGCTCCAAAGCTCCAGGGGCTCCCGAGGACCGGTGCTCG 240
Db 181 GGCCTCAAGCCCGACCGGCTCCAAAGCTCCAGGGGCTCCCGAGGACCGGTGCTCG 240
QY 241 GGCCTTCCTTCGGTCAGAAAGTCGCCCTCGGGGCGAGTTCTGCCAAAGGTTTCCTCG 300
Db 241 GGCCTTCCTTCGGTCAGAAAGTCGCCCTCGGGGCGAGTTCTGCCAAAGGTTTCCTCG 300
QY 301 AAAGAATCTGAGAGGGCGCAGTCTTGTAGCCGAGGAATCTCTGTGTAGCCCTTGAAGC 360
Db 301 AAAGAATCTGAGAGGGCGCAGTCTTGTAGCCGAGGAATCTCTGTGTAGCCCTTGAAGC 360
QY 361 CGCCAGCCCGCAGAAG 375
Db 361 CGCCAGCCCGCAGAAG 375

RESULT 3
HSA243396
LOCUS Homo sapiens mRNA for voltage-gated sodium channel beta-3 subunit (scn3b gene).
ACCESSION AJ243396
VERSION AJ243396.2 GI:7242612
KEYWORDS scn3b gene; voltage-gated sodium channel beta-3 subunit.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Morgan, K., Stevens, E.B., Shaw, B., Cox, P., Dixon, A.K., Lee, K., Pimock, R.D., Hughes, J., Richardson, P.J., Mizuguchi, K. and Jackson, A.P.
TITLE beta 3: an additional auxiliary subunit of the voltage-sensitive sodium channel that modulates channel gating with distinct kinetics
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (5), 2308-2313 (2000)
MEDLINE 20160948
PUBMED 10688874
REFERENCE 2
AUTHORS Morgan, K.
TITLE Direct Submission
JOURNAL Submitted (28-JUN-1999) Morgan K., Biochemistry, University of Cambridge, Tennis Court Road, Cambridge, CB2 1QW, UNITED KINGDOM
REMARK Revised by [4]
REFERENCE 3 (bases 1 to 1261)
AUTHORS Morgan, K.
TITLE Direct Submission
JOURNAL Submitted (13-MAR-2000) Morgan K., Biochemistry, University of Cambridge, Tennis Court Road, Cambridge, CB2 1QW, UNITED KINGDOM
COMMENT On Mar 14, 2000 this sequence version replaced gi:7160974.
FEATURES
    Location/Qualifiers
    1..1261
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /tissue_type="brain"
    1..1261
    /gene="scn3b"
    <1..375
    /gene="scn3b"
    376..1024
    /gene="scn3b"
    /function="modulates channel gating kinetics"
    /evidence=experimental
    /product="voltage-gated sodium channel beta-3 subunit"
    /protein_id="CAB76825.1"
    /db_xref="GI:7160975"
    /db_xref="GOA:Q9NY72"
    /db_xref="UniProt/Swiss-Prot:Q9NY72"
    /translation="MPAFNRLFPPLASLVLIVMSVCFVVPVSETEAVQGNPKMLR
    CISCMKREVEATVWFYRPEGGKDFLIYEYRNGHQEVESPFQGRLOWNGSKDLQD
    VITVLNVTLDNSGLYTCNVSREFEFAHRPFVKTTRLPLRVTGEAGDFTSVVSEI
    MMVILLVPLTLWLLIEMIYCYRKVSKAEAAQENASDYLAIPSENKENSAPVEE"
    1025..>1261
    /gene="scn3b"

Query Match 100.0%; Score 375; DB 9; Length 1261;
Best Local Similarity 100.0%; Pred. No. 1.4e-70;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCCTCCCTCCGAGTGTACCTTGGCGGCGAAACGAGCGAGCGCGCGAGTGG 60
Db 1 CCCTCCCTCCGAGTGTACCTTGGCGGCGAAACGAGCGAGCGCGCGAGTGG 60
QY 61 AAGCTGGAGTTCGGGGTGGCGGGAGCGGAGTGTCCGTGTGTGTGAGCGCGCGAGA 120
Db 61 AAGCTGGAGTTCGGGGTGGCGGGAGCGGAGTGTCCGTGTGTGTGAGCGCGCGAGA 120
QY 121 GCGGCGCGGAGCGGTGATCGGCTCCCTCGAACTGGGGAGGTCCAGTGGGGTCCGCTTAG 180
Db 121 GCGGCGCGGAGCGGTGATCGGCTCCCTCGAACTGGGGAGGTCCAGTGGGGTCCGCTTAG 180
```

```
Db 121 GCGGCGCGGAGGGCTGATCGCTCCCTCGAAGTGGGAGGTCCAGTGGGTGCTTAG 180
QY 181 GCGCCAAAGCCCAACCCGCTCCAAAGTCCAGGGCTCCCGAGGACCGGTCTCG 240
Db 181 GCGCCAAAGCCCAACCCGCTCCAAAGTCCAGGGCTCCCGAGGACCGGTGCTCG 240
QY 241 GCGCTTCCTTCGTCAGAAAGTCCGCCCTCGGGGCGAGTTCGTCGCAAGGGTTCTCTCG 300
Db 241 GCGCTTCCTTCGTCAGAAAGTCCGCCCTCGGGGCGAGTTCGTCGCAAGGGTTCTCTCG 300
QY 301 AAGAAATCTGAGAGGGCGAGTCTTGACCGAGGATCTCTGTGTAGCCTTGAAGC 360
Db 301 AAGAAATCTGAGAGGGCGAGTCTTGACCGAGGATCTCTGTGTAGCCTTGAAGC 360
QY 361 CGCAGCCCGCAGAG 375
Db 361 CGCAGCCCGCAGAG 375

RESULT 4
LOCUS AP000682/c
DEFINITION Homo sapiens chromosome 11 clone CMB9-32A1 map llq24, WORKING DRAFT
ACCESSION AP000682
VERSION AP000682.3 GI:9844967
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 127347)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Homo sapiens 127,347 Genomic DNA of llq24
JOURNAL Published Only in DataBase (1999)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (08-NOV-1999) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-42-778-9923, Fax:81-42-778-9924)
COMMENT On Aug 18, 2000 this sequence version replaced gi:8118870.
----- Genome Center
Center: RIKEN Genomic Sciences Center (GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
----- Project Information
Center project name: HumDraft11
Center clone name: CMB9-32A1
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 124961 bases at least Q40
Consensus quality: 125815 bases at least Q30
Consensus quality: 126286 bases at least Q20
Insert size: 126647; sum-of-contigs
Quality coverage: 12.71x in Q20 bases; sum-of-contigs
-----
NOTE: This is a 'working draft' sequence. It currently consists of
8 contigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs N, but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved
1 33532 contig of 33532 bp in length
33633 59223 contig of 25591 bp in length
```

```
59324 83007 contig of 23684 bp in length
83108 100196 contig of 17089 bp in length
100297 112313 contig of 12017 bp in length
112414 121043 contig of 8630 bp in length
121144 125917 contig of 4774 bp in length
126018 127347 contig of 1330 bp in length.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 33532: contig of 33532 bp in length
33633 33632: gap of 100 bp
33633 59223: contig of 25591 bp in length
59224 59223: gap of 100 bp
59324 83007: contig of 23684 bp in length
83108 83107: gap of 100 bp
83108 100196: contig of 17089 bp in length
100197 100296: gap of 100 bp
100297 112313: contig of 12017 bp in length
112314 112413: gap of 100 bp
112414 121043: contig of 8630 bp in length
121044 121143: gap of 100 bp
121144 125917: contig of 4774 bp in length
125918 126017: gap of 100 bp
126018 127347: contig of 1330 bp in length.

FEATURES
Location/Qualifiers
source
1..127347
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosomes="11"
/map="llq24"
/clone="CMB9-32A1"
misc_feature 1..33532
/notes="assembly_fragment"
misc_feature 33633..59223
/notes="assembly_fragment"
misc_feature 59324..83007
/notes="assembly_fragment"
misc_feature 83108..100196
/notes="assembly_fragment clone_end:SP6 vector_side:left"
misc_feature 100297..112313
/notes="assembly_fragment"
misc_feature 112414..121043
/notes="assembly_fragment clone_end:T7 vector_side:right"
misc_feature 121144..125917
/notes="assembly_fragment"
misc_feature 126018..127347
/notes="assembly_fragment"

ORIGIN
Query Match 99.5%; Score 373; DB 2; Length 127347;
Best Local Similarity 100.0%; Pred. No. 1.9e-70;
Matches 373; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 CTCCTTCGAGCTGAGCTTACCTCGGCGCAAAACGAGGAGCGGCGCGAGTGGAA 62
Db 29132 CTCCTTCGAGCTGAGCTTACCTCGGCGCAAAACGAGGAGCGGCGCGAGTGGAA 29073
QY 63 GCTGGAGTTCCGGGTGGGCGGCGGACGTGCTCGTGGTGTAGCGCGCGGAGAGC 122
Db 29072 GCTGGAGTTCCGGGTGGGCGGCGGAGCGGACGTGCTCGTGGTGTAGCGCGGAGAGC 29013
QY 123 GGGCGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGAGAGTCCAGTGGGTTCGCTTAGGG 182
Db 29012 GGGCGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGAGAGTCCAGTGGGTTCGCTTAGGG 28953
QY 183 CCAAAAGCCCCCAACCGCGGTCCAAAAGTCCCAAGGCGCTCCCGAGGACCGGTGCTCGGC 242
```

Db 28952 CCCAAGCCCCCACCAGCGGTCCAAAGCTCCAGAGCGCTCCAGCAGCCGCGTCTCGGC 28893
 QY 243 CTTCTCTCGGTGAGAAAGTCGCCCCCTCGGGGCGAGTTCGTGCTCCAAAGGGTTTCTCGAA 302
 Db 28892 CTTCTCTCGGTGAGAAAGTCGCCCCCTCGGGGCGAGTTCGTGCTCCAAAGGGTTTCTCGAA 28833
 QY 303 AGAATCTGAGAGGGCGCAGTCTTGACCGAGGGAATCTCTGTAGCTTGGAGCGG 362
 Db 28832 AGAATCTGAGAGGGCGCAGTCTTGACCGAGGGAATCTCTGTAGCTTGGAGCGG 28773
 QY 363 CCAGCCCCCAGAAG 375
 Db 28772 CCAGCCCCCAGAAG 28760

 RESULT 5
 AC063921/c
 LOCUS
 DEFINITION Homo sapiens chromosome 3 clone RP11-142P10, WORKING DRAFT
 AC063921
 AC063921.18 GI:20335587
 HTG: HTGS_PHASE1; HTGS_DRAFT.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 144833)
 Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
 Alebrooks,S.L., Amarantunge,H.C., Are,J.R., Avele,M., Banks,T.,
 Barbara,J., Benton,J., Binage,K., Blankenburg,K., Bonnin,D.,
 Bouay,J., Bowls,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
 Buhaq,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
 Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
 Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
 Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
 Douthwaite,K.J., Draper,D., Dugan-Rocha,S., Durbin,K.J.,
 Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
 Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
 Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
 Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
 Harris,C., Harris,K., Hart,M., Haviak,P., Hawes,A., Hernandez,J.,
 Hernandez,O., Hodgson,A., Hogue,M., Holloway,C., Hollins,B.,
 Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
 Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
 Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
 Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
 Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
 Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
 Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
 Massey,E., Mahoney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
 Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
 Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
 Nguyen,N., Nickerson,E., Nwokkwo,S., Ogih,M., Okunolu,G.,
 Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,I.,
 Peters,L., Pockens,A., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
 Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,
 Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
 Sodergren,E., Sonaite,T., Sparks,A., Stanley,H., Stone,H.,
 Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
 Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
 Usmani,K., Vaequez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
 Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
 Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
 Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
 Weinstein,G., and Gibbs,R.
 Direct Submission
 Unpublished
 2 (bases 1 to 144833)
 Worley,K.C.
 Direct Submission

JOURNAL

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

Submitted (22-APR-2000) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 144833)
 Worley,K.C.
 Direct Submission
 Submitted (08-JAN-2003) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Apr 28, 2002 this sequence version replaced gi:16117924.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: HAXY

Center clone name: RP11-142P10

----- Summary Statistics

Sequencing vector: M13;

Chemistry: Dye-Primer Bodipy: 14% of reads

Assembly: Dye-terminator Big Dye: 86% of reads

Assembly program: Phrap: version 0.990329

Consensus quality: 127260 bases at least Q40

Consensus quality: 132735 bases at least Q30

Consensus quality: 136183 bases at least Q20

Estimated insert size: 139078; sum-of-contigs estimation

Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 consists of 13 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence.
 * as soon as it is available and the accession number will
 be preserved.

1 2437: contig of 2437 bp in length
 2438 2537: gap of unknown length
 2538 4812: contig of 2275 bp in length
 4813 4912: gap of unknown length
 4913 8488: contig of 3576 bp in length
 8489 8588: gap of unknown length
 8589 11163: contig of 2575 bp in length
 11164 16638: contig of 5375 bp in length
 16639 16738: gap of unknown length
 16739 21870: contig of 5132 bp in length
 21871 21970: gap of unknown length
 21971 30780: contig of 8810 bp in length
 30781 30880: gap of unknown length
 30881 39267: contig of 8387 bp in length
 39268 39367: gap of unknown length
 39368 46890: contig of 7523 bp in length
 46891 46990: gap of unknown length
 46991 59656: contig of 12666 bp in length
 59657 59756: gap of unknown length
 59757 83309: contig of 23553 bp in length
 83310 83409: gap of unknown length
 83410 111857: contig of 28448 bp in length
 111858 111957: gap of unknown length
 111958 144833: contig of 32876 bp in length.

Location/Qualifiers

1..144833
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="3"
 /clone="RP11-142P10"

FEATURES
source

ORIGIN

TITLE
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 144833)
Worley,K.C.
AUTHORS
TITLE
Direct Submission


```
Query Match 99.5%; Score 373; DB 2; Length 144833;
Best Local Similarity 100.0%; Pred. No. 1.8e-70;
Matches 373; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CTCCCTCCGAGCTGAGCTTACCTCGGGCGCAACGAGCAGGCGGGCGCGAGTGGAA 62
Db 50353 CTCCCTCCGAGCTGAGCTTACCTCGGGCGCAACGAGCAGGCGGGCGCGAGTGGAA 50294

QY 63 GCTGGAGTTCGGGGTGGGGGAGCGGAGCTGTCGCTGTGCTGACGCCGCGGAGAGC 122
Db 50293 GCTGGAGTTCGGGGTGGGGGAGCGGAGCTGTCGCTGTGCTGACGCCGCGGAGAGC 50234

QY 123 GGGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGAGGTCCAGTGGGTGCGCTTAGGG 182
Db 50233 GGGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGAGGTCCAGTGGGTGCGCTTAGGG 50174

QY 183 CCCAAGCCCCCGGCTCCAAAGCTCCAGGGCTCCCGAGCAGCAGCGTGTCTCGGC 242
Db 50173 CCCAAGCCCCCGGCTCCAAAGCTCCAGGGCTCCCGAGCAGCAGCGTGTCTCGGC 50114

QY 243 CTTCTCTTCGCTCAGAAAGTCGCCCCCTGGGGCGAGTTCGTCCTCAAGGGTTTCCTCGAA 302
Db 50113 CTTCTCTTCGCTCAGAAAGTCGCCCCCTGGGGCGAGTTCGTCCTCAAGGGTTTCCTCGAA 50054

QY 303 AGAATCTGAGAGCGCGCAGTCTTGACCCGAGGGAATCTCTGTGTAGCTTGGAAAGCCG 362
Db 50053 AGAATCTGAGAGCGCGCAGTCTTGACCCGAGGGAATCTCTGTGTAGCTTGGAAAGCCG 49994

QY 363 CCAGCCCCCAGAAG 375
Db 49993 CCAGCCCCCAGAAG 49981

RESULT 6
AC021981/c 149800 bp DNA linear HTG 19-JUL-2000
LOCUS Homo sapiens chromosome 11 clone RP11-11C15 map 11, WORKING DRAFT
DEFINITION SEQUENCE, 29 unordered pieces.
ACCESSION AC021981
VERSION 1
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 149800)
AUTHORS Birren, B., Linton, L., Nussbaum, C., and Lander, E.
TITLE Homo sapiens chromosome 11, clone RP11-11C15
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 149800)
AUTHORS Birren, B., Linton, L., Nussbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Beckerly, R., Beda, F.,
Boguslavskiy, L., Boukhalter, B., Brown, A., Burkett, G., Castle, A.,
Choepey, Y., Collange, M., Collins, S., Collymore, A., Cooke, P.,
DeArellano, K., Dewar, K., Domino, M., Doyle, M., Fenestor, J.,
Ferreira, P., FitzHugh, W., Forrest, C., Gage, D., Galagan, J.,
Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
Landers, T., Lehoczy, J., Levine, R., Lieu, C., Liu, G., Locke, K.,
Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K.,
McPheeters, R., Meldrum, J., Meneus, L., Morrow, J., Naylor, J.,
Norman, C. H., O'Connor, T., O'Donnell, P., Olivari, T. M., Peterson, K.,
Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,
Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J.,
Zimmer, A., and Zody, M.
DIRECT SUBMISSION
JOURNAL Submitted (23-JAN-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On Jul 19, 2000 this sequence version replaced gi:6731265.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
```

http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR

Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information

Center project name: L1335
Center clone name: 11_C.15

----- Summary Statistics

Sequencing vector: M13; M7815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 135317 bases at least Q40
Consensus quality: 141872 bases at least Q30
Consensus quality: 144404 bases at least Q20
Insert size: 121000; agarose-fp
Insert size: 147000; sum-of-contigs
Quality coverage: 4.2 in Q20 bases; agarose-fp
Quality coverage: 3.5 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 29 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.

1
* 1441: contig of 1441 bp in length
* 1442: gap of 100 bp
* 1542: contig of 1342 bp in length
* 2884: gap of 100 bp
* 2884: contig of 1666 bp in length
* 4649: gap of 100 bp
* 4749: contig of 1363 bp in length
* 6112: gap of 100 bp
* 6212: contig of 1201 bp in length
* 7413: gap of 100 bp
* 7514: contig of 1537 bp in length
* 9051: gap of 100 bp
* 9150: contig of 1794 bp in length
* 10944: gap of 100 bp
* 11045: contig of 2159 bp in length
* 13204: gap of 100 bp
* 13304: contig of 2815 bp in length
* 16118: gap of 100 bp
* 16218: contig of 2455 bp in length
* 18673: gap of 100 bp
* 18774: contig of 2407 bp in length
* 21181: gap of 100 bp
* 21280: contig of 1920 bp in length
* 23201: gap of 100 bp
* 23301: contig of 3226 bp in length
* 26526: gap of 100 bp
* 26527: contig of 3029 bp in length
* 29655: gap of 100 bp
* 29755: contig of 4983 bp in length
* 34738: gap of 100 bp
* 34739: contig of 4556 bp in length
* 34839: gap of 100 bp
* 39395: contig of 4175 bp in length
* 39495: gap of 100 bp
* 43670: contig of 6179 bp in length
* 43770: gap of 100 bp
* 49949: contig of 5530 bp in length
* 50049: gap of 100 bp
* 55579: contig of 5281 bp in length
* 60959: gap of 100 bp
* 61060: contig of 5335 bp in length
* 66395: gap of 100 bp
* 66495: contig of 6453 bp in length
* 72948: gap of 100 bp

```
* 73048 80874: contig of 7827 bp in length
* 80875 80974: gap of 100 bp
* 80975 87847: contig of 6873 bp in length
* 87848 87947: gap of 100 bp
* 87948 97279: contig of 9332 bp in length
* 97280 97379: gap of 100 bp
* 97380 106037: contig of 8638 bp in length
* 106038 106137: gap of 100 bp
* 106138 114206: contig of 8069 bp in length
* 114207 114306: gap of 100 bp
* 114307 129830: contig of 15524 bp in length
* 129831 129830: gap of 100 bp
* 129831 149800: contig of 19870 bp in length.
```

FEATURES

source

```
1..149800
  Location/Qualifiers
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="11"
    /map="11"
    /clone="RP11-11C15"
    /clone_lib="RPC1-11 Human Male BAC"
    1..141
      /note="assembly_fragment"
    1542..2883
      /note="assembly_fragment"
    2984..4649
      /note="assembly_fragment"
    4750..6112
      /note="assembly_fragment"
    6213..7413
      /note="assembly_fragment"
    7514..9050
      /note="assembly_fragment"
    9151..10944
      /note="assembly_fragment"
    11045..13203
      /note="assembly_fragment"
    13304..16118
      /note="assembly_fragment"
    16219..18673
      /note="assembly_fragment"
    clone_end:r7
    vector_side:right
```

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

```
/note="assembly_fragment"
87948..97279
/note="assembly_fragment"
97380..106037
/note="assembly_fragment"
106138..114206
/note="assembly_fragment"
114307..129830
/note="assembly_fragment"
129831..149800
/note="assembly_fragment"
```

ORIGIN

```
Query Match      99.5%; Score 373; DB 2; Length 149800;
Best Local Similarity 100.0%; Pred.No.1.8e-70;
Matches 373; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CTCCTTCGAGCTGAGCTTACCTGGGCGCAACGAGCGAGCGAGGGGCGGAGTGGAA 62
DB 19234 CTCCTTCGAGCTGAGCTTACCTGGGCGCAACGAGCGAGCGAGGGGCGGAGTGGAA 19175

QY 63 GCTGAGTTCCGGGCTGGGCGGAGGCGACTGTCCGTGGTCTGAGCGCCCGGAGAGC 122
DB 19174 GCTGAGTTCCGGGCTGGGCGGAGGCGACTGTCCGTGGTCTGAGCGCCCGGAGAGC 19115

QY 123 GGGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGAGGTCAGTGGGTCGCTTAGGG 182
DB 19114 GGGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGAGGTCAGTGGGTCGCTTAGGG 19055

QY 183 CCCAAGCCCCCACCACCGGCTCCAAAGCTCCACAGGCTCCCGAGGCTCCCGAGCGCCGTCGCGC 242
DB 19054 CCCAAGCCCCCACCACCGGCTCCAAAGCTCCACAGGCTCCCGAGGCTCCCGAGCGCCGTCGCGC 18995

QY 243 CCTTCCTTCGGTCAAGAAAGTCGCCCCCTGGGGGCGAGTTGTCGCCAAAGGTTTCCTCGAA 302
DB 18994 CCTTCCTTCGGTCAAGAAAGTCGCCCCCTGGGGGCGAGTTGTCGCCAAAGGTTTCCTCGAA 18935

QY 303 AGAATCTGAGAGGGCGGAGTCTTGACCGAGGGAATCTCTGTGTAGCTTGGAGCGG 362
DB 18934 AGAATCTGAGAGGGCGGAGTCTTGACCGAGGGAATCTCTGTGTAGCTTGGAGCGG 18875

QY 363 CCAGCCCCCAGAAG 375
DB 18874 CCAGCCCCCAGAAG 18862
```

RESULT 7

AC024604

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

```
172546 bp DNA linear HTG 30-AUG-2001
Homo sapiens chromosome Chromosome 10 clone RP11-331G19, WORKING
DRAFT SEQUENCE, 22 unordered pieces.
AC024604
AC024604.3 GI:8389428
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_CANCELLED.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Smith,D.R.
Genome Therapeutics Corporation Sequencing Center: Human Genome
Sequence Data
Unpublished
2 (bases 1 to 172546)
Smith,D.R.
Direct Submission
Submitted (01-MAR-2000) Genome Therapeutics Corporation, 100 Beaver
Street, Waltham, MA 02453, USA
On Jun 9, 2000 this sequence version replaced gi:7549605.
----- Genome Center
Center: Genome Therapeutics Corporation
Center code: GTC
Web site: http://www.genomecorp.com/
```

Contact: gtc-seqcenter@genomecorp.com
 ----- Project Information
 Center project name: hg202
 ----- Summary Statistics

Sequencing vector: N/A
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 990315
 Consensus quality: 152033 bases at least Q40
 Consensus quality: 163535 bases at least Q30
 Consensus quality: 165256 bases at least Q20
 Insert size: 170446; sum-of-contigs
 Quality coverage: 4.3x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 22 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 1339: contig of 1339 bp in length
 * 1340 1439: gap of unknown length
 * 1440 2789: contig of 1350 bp in length
 * 2790 2889: gap of unknown length
 * 2890 4562: contig of 1673 bp in length
 * 4563 4662: gap of unknown length
 * 4663 6344: contig of 1682 bp in length
 * 6345 6444: gap of unknown length
 * 6445 9221: contig of 2777 bp in length
 * 9222 9321: gap of unknown length
 * 9323 11997: contig of 2676 bp in length
 * 11998 12097: gap of unknown length
 * 12098 14867: contig of 2770 bp in length
 * 14868 14967: gap of unknown length
 * 14968 19919: contig of 4952 bp in length
 * 19920 20019: gap of unknown length
 * 20020 24125: contig of 4106 bp in length
 * 24126 24225: gap of unknown length
 * 24226 29444: contig of 5219 bp in length
 * 29445 29544: gap of unknown length
 * 29545 34568: contig of 5024 bp in length
 * 34569 34668: gap of unknown length
 * 34669 39414: contig of 4746 bp in length
 * 39415 39514: gap of unknown length
 * 39515 45885: contig of 6371 bp in length
 * 45886 45985: gap of unknown length
 * 45986 50997: contig of 5012 bp in length
 * 50998 51097: gap of unknown length
 * 51098 59143: contig of 8046 bp in length
 * 59144 59243: gap of unknown length
 * 59244 67779: contig of 8536 bp in length
 * 67780 67879: gap of unknown length
 * 67880 76429: contig of 8550 bp in length
 * 76430 76529: gap of unknown length
 * 76530 86704: contig of 10175 bp in length
 * 86705 86804: gap of unknown length
 * 86805 103147: contig of 16343 bp in length
 * 103148 103247: gap of unknown length
 * 103248 116343: contig of 13096 bp in length
 * 116344 116443: gap of unknown length
 * 116444 135318: contig of 18875 bp in length
 * 135319 135419: gap of unknown length
 * 135419 172546: contig of 37128 bp in length.

FEATURES
 source

1. .172546
 /location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="Chromosome 10"
 /clone_lib="RPCI-11"

ORIGIN

Query Match 99.5%; Score 373; DB 9; Length 178169;
 Best Local Similarity 100.0%; Pred. No. 1.8e-70;

Query Match 99.5%; Score 373; DB 2; Length 172546;
 Best Local Similarity 100.0%; Pred. No. 1.8e-70;
 Matches 373; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 CTCCTCTCCGAGCTGAGCTTACCTGGCGCGCAACACGAGCGAGCGCGCGCGAGTGGAA 62
 DB 123777 CTCCCTTCCGAGCTGAGCTTACCTGGCGCGCAACACGAGCGAGCGCGCGAGTGGAA 122436
 QY 63 GCTGGAGTTCCGGGGTGGCGGGAGGCGACTGCTCCGTGTGTGAGCGCGCGCGAGAGC 122
 DB 122437 GCTGGAGTTCCGGGGTGGCGGGAGGCGACTGCTCCGTGTGTGAGCGCGCGCGAGAGC 122496
 QY 123 GGGCGGGAGCGGCTGATCCGGCTCCCTCGAATCTGGGAGGTCCAGTGGGGTCCGCTTAGGG 182
 DB 122497 GGGCGGGAGCGGCTGATCCGGCTCCCTCGAATCTGGGAGGTCCAGTGGGGTCCGCTTAGGG 122556
 QY 183 CCCAAAGCCCCACCCCGGCTCCAAAAGCTCCAGAGGCTCCCCAGGACCGCGTCTCGGC 242
 DB 122557 CCCAAAGCCCCACCCCGGCTCCAAAAGCTCCAGAGGCTCCCCAGGACCGCGTCTCGGC 122616
 QY 243 CCTTCTCTCGGTGAGAAAGTCGCCCTCTGGGGGAGTTCGTCCCAAAGGGTTCTCTCGAA 302
 DB 122617 CCTTCTCTCGGTGAGAAAGTCGCCCTCTGGGGGAGTTCGTCCCAAAGGGTTCTCTCGAA 122676
 QY 303 AGAATCTGAGAGGCGCGCAGTCTTTCACCGAGGGAATCTCTCTGTGTAGCCCTTGAAGCCG 362
 DB 122677 AGAATCTGAGAGGCGCGCAGTCTTTCACCGAGGGAATCTCTCTGTGTAGCCCTTGAAGCCG 122736
 QY 363 CCAGCCCCAGAG 375
 DB 122737 CCAGCCCCAGAG 122749

RESULT 8
 AP002765/c

LOCUS Homo sapiens genomic DNA, chromosome 11q clone:RP11-634B22,
 DEFINITION complete sequence.
 ACCESSION AP002765 GI:16751488
 VERSION AP002765.3
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 AUTHORS

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 Homo sapiens genomic DNA
 Published Only in Database (2000)
 REFERENCE 2 (bases 1 to 178169)
 AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 Direct Submission

TITLE
 JOURNAL

Submitted (10-JUL-2000) Masahira Hattori, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
 1-7-22, Suehiro-chou, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
 Tel:81-45-503-9111, Fax:81-45-503-9170)
 On Nov 5, 2001 this sequence version replaced gi:12381934.

COMMENT
 FEATURES

source
 1. .178169
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11q"
 /clone="RP11-634B22"

ORIGIN

```

Matches 373; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 CTCCTTCCGAGCTGAGCTTACCTGGGCGCAAAACGAGCGAGCGAGGCGCGAGTGGAA 62
Db 107533 CTCCTTCCGAGCTGAGCTTACCTGGGCGCAAAACGAGCGAGCGAGGCGCGAGTGGAA 107474
QY 63 GCTGGAGTTCCCGGGTGGGCGGGAGGCGACTGTCCGTGGTGTCTGAGCCCGCGAGAGC 122
Db 107473 GCTGGAGTTCCCGGGTGGGCGGGAGGCGACTGTCCGTGGTGTCTGAGCCCGCGAGAGC 107414
QY 123 GGGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGGAGGTCAGTGGGGTGCCTTAGGG 182
Db 107413 GGGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGGAGGTCAGTGGGGTGCCTTAGGG 107354
QY 183 CCCAAGCCCCCACCAGCGGCTCCAAAGCTCCAGGCGCTCCCGAGCACCAGGTCCTCGGC 242
Db 107353 CCCAAGCCCCCACCAGCGGCTCCAAAGCTCCCGAGGCGCTCCCGAGCACCAGGTCCTCGGC 107294
QY 243 CTTCTCTCGGTCAGAAAGTCGCGCCCTGGGGGCGAGTTCGTCGAAAGGGTTTCCTCGAA 302
Db 107293 CTTCTCTCGGTCAGAAAGTCGCGCCCTGGGGGCGAGTTCGTCGAAAGGGTTTCCTCGAA 107234
QY 303 AGAATCTGAGAGGCGGCGAGTCTTGACCGAGGGAATCTCTGTGTAGCTTGGAGCGG 362
Db 107233 AGAATCTGAGAGGCGGCGAGTCTTGACCGAGGGAATCTCTGTGTAGCTTGGAGCGG 107174
QY 363 CCAGCCCCCAGAAG 375
Db 107173 CCAGCCCCCAGAAG 107161

```

```

RESULT 9
AC069539/c
LOCUS AC069539 181471 bp DNA linear PRI 05-AUG-2004
DEFINITION Homo sapiens chromosome 11 clone RP11-321E15, complete sequence.
ACCESSION AC069539
VERSION AC069539.5 GI:20270093
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 181471)
AUTHORS Smith,D.R.
TITLE Genome Therapeutics Corporation Sequencing Center: Human Genome
SEQUENCE DATA
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 181471)
AUTHORS Smith,D.R.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-2000) Genome Therapeutics Corporation, 100 Beaver
Street, Waltham, MA 02453, USA
REFERENCE 3 (bases 1 to 181471)
AUTHORS Smith,D.R.
TITLE Direct Submission
JOURNAL Submitted (19-OCT-2000) Genome Therapeutics Corporation, 100 Beaver
Street, Waltham, MA 02453, USA
REFERENCE 4 (bases 1 to 181471)
AUTHORS Smith,D.R.
TITLE Direct Submission
JOURNAL Submitted (23-APR-2002) Genome Therapeutics Corporation, 100 Beaver
Street, Waltham, MA 02453, USA
COMMENT On Apr 23, 2002 this sequence version replaced gi:10881056.
FEATURES
source
1. .181471
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosomes="11"
/clone="RP11-321E15"
/clone_lib="RPC1-11"
ORIGIN

```

```

Query Match 99.5%; Score 373; DB 9; Length 181471;
Best Local Similarity 100.0%; Pred. No. 1.8e-70;
Matches 373; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 CTCCTTCCGAGCTGAGCTTACCTGGGCGCAAAACGAGCGAGCGAGGCGCGAGTGGAA 62
Db 79671 CTCCTTCCGAGCTGAGCTTACCTGGGCGCAAAACGAGCGAGCGAGGCGCGAGTGGAA 79612
QY 63 GCTGGAGTTCCCGGGTGGGCGGGAGGCGACTGTCCGTGGTGTCTGAGCCCGCGAGAGC 122
Db 79611 GCTGGAGTTCCCGGGTGGGCGGGAGGCGACTGTCCGTGGTGTCTGAGCCCGCGAGAGC 79552
QY 123 GGGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGGAGGTCAGTGGGGTGCCTTAGGG 182
Db 79551 GGGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGGAGGTCAGTGGGGTGCCTTAGGG 79492
QY 183 CCCAAGCCCCCACCAGCGGCTCCAAAGCTCCAGGCGCTCCCGAGCACCAGGTCCTCGGC 242
Db 79491 CCCAAGCCCCCACCAGCGGCTCCAAAGCTCCCGAGGCGCTCCCGAGCACCAGGTCCTCGGC 79432
QY 243 CTTCTCTCGGTCAGAAAGTCGCGCCCTGGGGGCGAGTTCGTCGAAAGGGTTTCCTCGAA 302
Db 79431 CTTCTCTCGGTCAGAAAGTCGCGCCCTGGGGGCGAGTTCGTCGAAAGGGTTTCCTCGAA 79372
QY 303 AGAATCTGAGAGGCGGCGAGTCTTGACCGAGGGAATCTCTGTGTAGCTTGGAGCGG 362
Db 79371 AGAATCTGAGAGGCGGCGAGTCTTGACCGAGGGAATCTCTGTGTAGCTTGGAGCGG 79312
QY 363 CCAGCCCCCAGAAG 375
Db 79311 CCAGCCCCCAGAAG 79299

```

AP002749 178431 bp DNA linear HTG 13-JUL-2000
Homo sapiens chromosome 11 clone RP11-158J21 map 11q24, WORKING
DRAFT SEQUENCE, 41 unordered pieces.

AP002749
AP002749.1 GI:9188598
HTG; HTGS PHASE1; HTGS_DRAFT.
Homo sapiens (human)
Homo sapiens
Homo sapiens
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 178431)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Homo sapiens 178,431 genomic DNA of 11q24
Published Only in DataBase (2000)
2 (bases 1 to 178431)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (10-JUL-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-42-778-9923, Fax:81-42-778-9924)
----- Genome Center
Center: RIKEN Genomic Sciences Center(GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
----- Project Information
Center Project name: HumDraft11
Center clone name: RP11-158J21
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 155982 bases at least Q40
Consensus quality: 165758 bases at least Q30

Consensus quality: 170239 bases at least Q20
 Insert size: 174431; sum-of-contigs
 Quality coverage: 4.20x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 41 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

```

1 13012 contig of 13012 bp in length
13113 24677 contig of 11565 bp in length
24778 34640 contig of 8863 bp in length
34741 43196 contig of 8456 bp in length
43297 50126 contig of 6830 bp in length
50227 57725 contig of 7499 bp in length
57826 64612 contig of 6287 bp in length
64713 70999 contig of 6287 bp in length
71100 76188 contig of 5089 bp in length
76289 83580 contig of 7292 bp in length
83681 90511 contig of 6831 bp in length
90612 95493 contig of 4882 bp in length
95594 102262 contig of 6669 bp in length
102363 105866 contig of 3504 bp in length
105967 111142 contig of 5176 bp in length
111243 115923 contig of 4681 bp in length
116024 121747 contig of 5724 bp in length
121848 125532 contig of 3685 bp in length
125633 127927 contig of 2295 bp in length
127928 131184 contig of 3157 bp in length
131285 132725 contig of 1441 bp in length
132826 135566 contig of 2741 bp in length
135667 138000 contig of 2334 bp in length
138001 141194 contig of 3094 bp in length
141295 143619 contig of 2325 bp in length
143720 146448 contig of 2729 bp in length
146549 152157 contig of 2983 bp in length
152258 155623 contig of 3366 bp in length
155724 158335 contig of 2611 bp in length
158435 160509 contig of 2075 bp in length
160610 162343 contig of 1734 bp in length
162444 164809 contig of 2366 bp in length
164910 169169 contig of 2033 bp in length
169170 171052 contig of 1783 bp in length
171053 172933 contig of 1781 bp in length
172934 174181 contig of 1147 bp in length
174280 175729 contig of 1449 bp in length
175730 176905 contig of 1076 bp in length
176906 177006 contig of 1426 bp in length.

1 13012: contig of 13012 bp in length
13113 13112: gap of 100 bp
24677 24677: contig of 11565 bp in length
24778 24777: gap of 100 bp
34640 34640: contig of 8863 bp in length
34741 34740: gap of 100 bp
43196 43196: contig of 8456 bp in length
43297 43296: gap of 100 bp
50126 50126: contig of 6830 bp in length
50227 50226: gap of 100 bp
57725 57725: contig of 7499 bp in length
57826 57825: gap of 100 bp
64612: contig of 6287 bp in length

```

NOTE: This is a 'working draft' sequence. It currently consists of 41 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```

1 13012: contig of 13012 bp in length
13113 13112: gap of 100 bp
24677 24677: contig of 11565 bp in length
24778 24777: gap of 100 bp
34640 34640: contig of 8863 bp in length
34741 34740: gap of 100 bp
43196 43196: contig of 8456 bp in length
43297 43296: gap of 100 bp
50126 50126: contig of 6830 bp in length
50227 50226: gap of 100 bp
57725 57725: contig of 7499 bp in length
57826 57825: gap of 100 bp
64612: contig of 6287 bp in length

```

FEATURES

source

Location/Qualifiers
 1..178431
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

```

* 64613 64712: gap of 100 bp
* 64713 70999: contig of 6287 bp in length
* 71000 71099: gap of 100 bp
* 71100 76188: contig of 5089 bp in length
* 76189 76288: gap of 100 bp
* 76289 83580: contig of 7292 bp in length
* 83581 83680: gap of 100 bp
* 83681 90511: contig of 6831 bp in length
* 90512 90611: gap of 100 bp
* 90612 95493: contig of 4882 bp in length
* 95494 95594: gap of 100 bp
* 95594 102262: contig of 6669 bp in length
* 102263 102362: gap of 100 bp
* 102363 105866: contig of 3504 bp in length
* 105867 105966: gap of 100 bp
* 105967 111142: contig of 5176 bp in length
* 111143 111242: gap of 100 bp
* 111243 115923: contig of 4681 bp in length
* 115924 116024: gap of 100 bp
* 116024 121747: contig of 5724 bp in length
* 121748 121847: gap of 100 bp
* 121848 125532: contig of 3685 bp in length
* 125533 125632: gap of 100 bp
* 125633 127927: contig of 2295 bp in length
* 127928 128027: gap of 100 bp
* 128028 131184: contig of 3157 bp in length
* 131185 131284: gap of 100 bp
* 131285 132725: contig of 1441 bp in length
* 132726 132825: gap of 100 bp
* 132826 135566: contig of 2741 bp in length
* 135567 135666: gap of 100 bp
* 135667 138000: contig of 2334 bp in length
* 138001 138100: gap of 100 bp
* 138101 141194: contig of 3094 bp in length
* 141195 141294: gap of 100 bp
* 141295 143619: contig of 2325 bp in length
* 143620 143719: gap of 100 bp
* 143720 146448: contig of 2729 bp in length
* 146449 146548: gap of 100 bp
* 146549 149531: contig of 2983 bp in length
* 149532 149631: gap of 100 bp
* 149632 152157: contig of 2526 bp in length
* 152158 152257: gap of 100 bp
* 152258 155623: contig of 3366 bp in length
* 155624 155723: gap of 100 bp
* 155724 158334: contig of 2611 bp in length
* 158335 158434: gap of 100 bp
* 158435 160509: contig of 2075 bp in length
* 160510 160609: gap of 100 bp
* 160610 162343: contig of 1734 bp in length
* 162344 162443: gap of 100 bp
* 162444 164809: contig of 2366 bp in length
* 164810 164909: gap of 100 bp
* 164910 167036: contig of 2127 bp in length
* 167037 167136: gap of 100 bp
* 167137 169169: contig of 2033 bp in length
* 169170 169269: gap of 100 bp
* 169270 171052: contig of 1783 bp in length
* 171053 171152: gap of 100 bp
* 171153 172933: contig of 1781 bp in length
* 172934 173033: gap of 100 bp
* 173034 174180: contig of 1147 bp in length
* 174181 174280: gap of 100 bp
* 174281 175729: contig of 1449 bp in length
* 175730 175829: gap of 100 bp
* 175830 176905: contig of 1076 bp in length
* 176906 177005: gap of 100 bp
* 177006 178431: contig of 1426 bp in length.

```

```
/chromosome="11"
/map="11q24"
/clone="RP11-158J21"
1. .13012
/note="assembly_fragment"
13113. .24677
/note="assembly_fragment"
24778. .34640
/note="assembly_fragment"
34741. .43196
/note="assembly_fragment"
43297. .50126
/note="assembly_fragment"
50227. .57725
/note="assembly_fragment"
57826. .64612

Query Match 93.3%; Score 350; DB 2; Length 178431;
Best Local Similarity 99.5%; Pred.No.1.6e-65;
Matches 372; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 3 CTCCTTCCGAGCTGAGCTTACCTTGGCGCAACGAGCGAGCGGGCGCGAGTGGAA 62
DB 111367 CTCCTTCCGAGCTGAGCTTACCTTGGCGCAACGAGCGAGCGCA-GGCGCGAGTGGAA 111425

QY 63 GCTGAGTTCGGGTGGCGGGAGGAGCTGTGCTGTGCTGAGCGCCGCGAGAGC 122
DB 111426 GCTGAGTTCGGGTGGCGGGAGGAGCTGTGCTGTGCTGAGCGCCGCGAGAGC 111485

QY 123 GGGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGAGGTCCAGTGGGTGCTTAGGG 182
DB 111486 GGGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGAGGTCCAGTGGGTGCTTAGGG 111545

QY 183 -CCCAAGCCCGGCTCCAAAGCTCCAGGGCTCCCGAGGCAACCGGTGCTCGG 241
DB 111546 CCCAAAGCCCGGCTCCAAAGCTCCCGAGGCTCCCGAGGCAACCGGTGCTCGG 111605

QY 242 CCTTCTTCGCTGAGAAAGTGCCTTGGGGGCGAGTTCGTCGCAAGGGTTCTCGA 301
DB 111606 CCTTCTTCGCTGAGAAAGTGCCTTGGGGGCGAGTTCGTCGCAAGGGTTCTCGA 111665

QY 302 AAGATCTCAGAGGGCGAGTCTTGACCGAGGAGTCTCTGTGTGATGCTTGAAGCC 361
DB 111666 AAGATCTCAGAGGGCGAGTCTTGACCGAGGAGTCTCTGTGTGATGCTTGAAGCC 111725

QY 362 GCCAGCCCGAGAAG 375
DB 111726 GCCAGCCCGAGAAG 111739

RESULT 11
AC145484/c
LOCUS
DEFINITION Lemur catta clone LB2-244B5, WORKING DRAFT SEQUENCE, 3 ordered
pieces.
AC145484
VERSION AC145484.1 GI:32880237
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT.
SOURCE Lemur catta (ring-tailed lemur)
ORGANISM Lemur catta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Strepsirhini; Lemuridae; Lemur.
1 (bases 1 to 174285)
Cheng,J.-F., Hamilton,M., Peng,Y., Mukherjee,S., Hosseini,R.,
Peng,Z., Malinov,I. and Rubin,E.M.
Direct Submission
Unpublished
2 (bases 1 to 174285)
Cheng,J.-F., Hamilton,M., Peng,Y., Mukherjee,S., Hosseini,R.,
Peng,Z., Malinov,I. and Rubin,E.M.
Direct Submission
Submitted (17-JUL-2003) Genome Sciences, Lawrence Berkeley National
Laboratory, 1 Cyclotron Rd., Berkeley, CA 94720, USA
```

COMMENT

Sequence Produced by Berkeley PGA
Web site: <http://pga.lbl.gov>
Center Code: PGABERK
Center Project Name: L133
Bac Clone Name: LB2-244B5

This sequence has been compared to sequences of other species
using Vista (<http://www-gsd.lbl.gov/VISTA>). The results can be
viewed at:
http://pga.lbl.gov/cgi-bin/search_cvcgd?type=n&value=ZNF202

The order-orientation of the draft sequence was accomplished by
using:
Avid (<http://baboon.math.berkeley.edu/mavid/>),
Lagan (<http://lagan.stanford.edu/>) and paired end information.

Funding agent: Programs for Genomic Applications (NHLBI)

If the Bac Library Name is LB1 to LB4, please see website
for the description: <http://www-gsd.lbl.gov/cheng/BAC.html>
These libraries are available through the BACPAC Resources Center:
<http://www.chori.org/bacpac/libraryres.htm> as LBNL-1 to LBNL-4.

Summary Statistics:

Sequencing vector: Plasmid; pUC18

Chemistry: Dye-terminator Big Dye

Assembly program: Phrap version 0.990329.

* NOTE: This is a 'working draft' sequence. It currently

* consists of 3 contigs. Gaps between the contigs

* are represented as runs of N. The order of the pieces

* is believed to be correct as given, however the sizes

* of the gaps between them are based on estimates that have

* provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

* 1 87364: contig of 87364 bp in length

* 87365 87464: gap of unknown length

* 87465 151123: contig of 63659 bp in length

* 151124 151223: gap of unknown length

* 151224 174285: contig of 23062 bp in length.

FEATURES

source

1..174285
/organism="Lemur catta"
/mol_type="genomic DNA"
/db_xref="taxon:9447"
/clone="LB2-244B5"

ORIGIN

| | | | | |
|-----------------------|-----------------|--|------------|----------------|
| Query Match | 43.7%; | Score 164; | DB 2; | Length 174285; |
| Best Local Similarity | 75.8%; | Pred. No. 1.5e-25; | | |
| Matches 288; | Conservative 0; | Mismatches 70; | Indels 22; | Gaps 6; |
| QY | 12 | GAGCTGAGCTTACCTTGGCGCAACGAGCGAGCGGGCGCGAGTGGAAAGCTGAGATT | 71 | |
| DB | 22530 | GAGCCGAGCTTCCCGCGCGCAACGAGAAAGGC-CGGGAGCAAGTGGGATTTGAAGTT | 22472 | |
| QY | 72 | CCGGGGTGGGGGAGGCGACTGTCCGTGGTGTGAGC-----GCCGG | 115 | |
| DB | 22471 | CC-GGGTGGGGGGAAGCCACTGTCTCTGTGTGTCGCGGATCCAGCAACCGGCTCC | 22413 | |
| QY | 116 | CGAGAGCGGGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGGAGTCCAGTGGGGTCG | 175 | |
| DB | 22412 | GGGGAGAGCGCGCAGAGCGCGTGTCTC-GCTCCCCGAAATGGGGCGCTCTGCTGGGAAG | 22354 | |
| QY | 176 | CTTAGGGCCCAAGCCCGCCACCGCGCTCCAAAGCTCCAGGGCTCCCGAGCAACCGGT | 235 | |
| DB | 22353 | CTCCGAG-CAGAGACCCCGCGCCCGCCACAGCCCGGAGCTCTCCCGAGCGGGA | 22296 | |
| QY | 236 | GCTCGGCCCTTCTTCGTCAGAAAGTCGCCCGCTGGGGGCGAGTTCGTCGCAAGGGTTT | 295 | |
| DB | 22295 | GGCTGGTCTTCTTTCAATCGAAAGTCGCCCGCTTGGGGCAGTTCGTCGTCGAGAGGGTTT | 22236 | |

```

QY      296  CCTCGAAGAATCTGAGAGGGCGCAGTCTTGTGACCGGGAATCTCTGTGTAGCCCTTG 355
      |||||||
Db      22235  CCTCGAAGAATCTGAGAGGGCGCAGTCTTGTGACCGGGAATCTCTGTGTAGCC-GG 22177

QY      356  GAAGCCCGCCGCCCGAAG 375
      |||||||
Db      22176  GAAGCCAACAGCTGAAGAAG 22157

RESULT 12
AC013796
LOCUS      Homo sapiens clone RP11-20H9, LOW-PASS SEQUENCE SAMPLING.
DEFINITION AC013796
ACCESSION AC013796
VERSION    AC013796.4 GI:12583842
KEYWORDS  HTG; HTGS_PHASE0.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
            Baldwin,J., Barna,N., Beckerly,R., Boguslavskiy,L., Boukhgalter,B.,
            Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
            Cooke,P., Dearellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
            Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
            Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Lehoczy,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
            McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
            Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
            Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
            Direct Submission
            Submitted (15-NOV-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
            On Jan 27, 2001 this sequence version replaced gi:7582586.
            All repeats were identified using RepeatMasker:
            Snit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            ----- Genome Center
            Center: Whitehead Institute/ MIT Center for Genome Research
            Center code: WtBR
            Web site: http://www-seq.wi.mit.edu
            Contact: sequence_submissions@genome.wi.mit.edu
            ----- Project Information
            Center project name: L3904
            Center clone name: 20_H_9
            -----
            * NOTE: This record contains 83 individual
            * sequencing reads that have not been assembled into
            * contigs. Runs of N are used to separate the reads
            * and the order in which they appear is completely
            * arbitrary. Low-pass sequence sampling is useful for
            * identifying clones that may be gene-rich and allows
            * overlap relationships among clones to be deduced.
            * However, it should not be assumed that this clone
            * will be sequenced to completion. In the event that
            * the record is updated, the accession number will
            * be preserved.
            * 1      683: contig of 683 bp in length
            *      684      783: gap of 100 bp
            *      784      1488: contig of 705 bp in length
            *      1488      1588: gap of 100 bp
            *      1589      2296: contig of 708 bp in length
            *      2297      2396: gap of 100 bp

            2397      3101: contig of 705 bp in length
            3101: gap of 100 bp
            3201: contig of 715 bp in length
            3916: gap of 100 bp
            4016: contig of 721 bp in length
            4837: contig of 100 bp
            5557: gap of 100 bp
            5657: contig of 715 bp in length
            6372: gap of 100 bp
            6472: contig of 683 bp in length
            7155: gap of 100 bp
            7255: contig of 702 bp in length
            7957: gap of 100 bp
            8057: contig of 716 bp in length
            8773: gap of 100 bp
            8873: contig of 704 bp in length
            9577: gap of 100 bp
            10392: contig of 715 bp in length
            10492: gap of 100 bp
            11179: contig of 687 bp in length
            11279: gap of 100 bp
            11984: contig of 705 bp in length
            12084: gap of 100 bp
            12791: contig of 707 bp in length
            12891: gap of 100 bp
            13600: contig of 709 bp in length
            13700: gap of 100 bp
            14409: contig of 709 bp in length
            14509: gap of 100 bp
            15221: contig of 712 bp in length
            15321: gap of 100 bp
            16025: contig of 704 bp in length
            16125: gap of 100 bp
            16822: contig of 697 bp in length
            16922: gap of 100 bp
            17621: contig of 699 bp in length
            17721: gap of 100 bp
            18423: contig of 702 bp in length
            18523: gap of 100 bp
            19241: contig of 718 bp in length
            19341: gap of 100 bp
            20056: contig of 715 bp in length
            20156: gap of 100 bp
            20870: contig of 714 bp in length
            20970: gap of 100 bp
            21671: contig of 701 bp in length
            21771: gap of 100 bp
            22480: contig of 709 bp in length
            22580: gap of 100 bp
            23293: contig of 713 bp in length
            23393: gap of 100 bp
            24095: contig of 702 bp in length
            24195: gap of 100 bp
            24916: contig of 721 bp in length
            25016: gap of 100 bp
            25726: contig of 710 bp in length
            25826: gap of 100 bp
            26534: contig of 708 bp in length
            26634: gap of 100 bp
            27333: contig of 699 bp in length
            27433: gap of 100 bp
            28131: contig of 698 bp in length
            28231: gap of 100 bp
            28926: contig of 695 bp in length
            29026: gap of 100 bp
            29758: contig of 732 bp in length
            29858: gap of 100 bp
            30566: contig of 708 bp in length
            30666: gap of 100 bp
            31368: contig of 702 bp in length
            31468: gap of 100 bp
            32183: contig of 715 bp in length

```

```
* 32184 32283: gap of 100 bp
* 32284 32981: contig of 698 bp in length
* 32982 33081: gap of 100 bp
* 33082 33798: contig of 717 bp in length
* 33799 33898: gap of 100 bp
* 33899 34634: contig of 736 bp in length
* 34635 34734: gap of 100 bp
* 34735 35433: contig of 699 bp in length
* 35434 35533: gap of 100 bp
* 35534 36253: contig of 720 bp in length
* 36254 36353: gap of 100 bp
* 36354 37079: contig of 726 bp in length
* 37080 37179: gap of 100 bp
* 37180 37893: contig of 714 bp in length
* 37894 37993: gap of 100 bp
* 37994 38700: contig of 707 bp in length
* 38701 38801: gap of 100 bp
* 38801 39505: contig of 705 bp in length
* 39506 39605: gap of 100 bp
* 39606 40316: contig of 711 bp in length
* 40317 40416: gap of 100 bp
* 40417 41131: contig of 715 bp in length
* 41132 41231: gap of 100 bp
* 41232 41939: contig of 708 bp in length
* 41940 42039: gap of 100 bp
* 42040 42748: contig of 709 bp in length
* 42749 42848: gap of 100 bp
* 42849 43555: contig of 707 bp in length
* 43556 43655: gap of 100 bp
* 43656 44368: contig of 713 bp in length
* 44369 44468: gap of 100 bp
* 44470 45271: contig of 703 bp in length
* 45172 45271: gap of 100 bp
* 45272 45943: contig of 672 bp in length
* 45944 46043: gap of 100 bp
* 46044 46749: contig of 706 bp in length
* 46750 46849: gap of 100 bp
* 46850 47551: contig of 701 bp in length
* 47551 47651: gap of 100 bp
* 47651 48350: contig of 700 bp in length
* 48351 48450: gap of 100 bp
* 48451 49148: contig of 698 bp in length
* 49149 49248: gap of 100 bp
* 49249 49361: contig of 713 bp in length
* 49362 50061: gap of 100 bp
* 50062 50751: contig of 690 bp in length
* 50752 50851: gap of 100 bp
* 50852 51567: contig of 716 bp in length
* 51568 51667: gap of 100 bp
* 51668 52400: contig of 733 bp in length
* 52401 52500: gap of 100 bp
* 52501 53225: contig of 725 bp in length
* 53226 53325: gap of 100 bp
* 53326 53974: contig of 649 bp in length
* 53975 54074: gap of 100 bp
* 54075 54781: contig of 707 bp in length
* 54782 54881: gap of 100 bp
* 54882 55608: contig of 727 bp in length
* 55609 55708: gap of 100 bp
* 55709 56426: contig of 718 bp in length
* 56427 56526: gap of 100 bp
* 56527 57244: contig of 718 bp in length
* 57245 57344: gap of 100 bp
* 57345 58051: contig of 707 bp in length
* 58052 58151: gap of 100 bp
* 58152 58860: contig of 709 bp in length

Query Match 32.5%; Score 122; DB 2; Length 66980;
Best Local Similarity 99.3%; Pred. No. 1.8e-16;
Matches 133; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 242 CCCTTCCTCGTCAGAAAGTCGCCCGGCGAGTTCGTCGCGAGGGTTCTCGA 301
|||||
```

```
Db 63886 CCCTTCCTCGTCAGAAAGTCGCCCGGCGAGTTCGTCGCC-AGGGTTTCTCGA 63944
QY 302 AAGAAATCTGAGAGGCGCGAGTCTTTGACCGAGGGAATCTCTGTGTAGCCTTGAAGCC 361
|||||
63945 AAGAATCTGAGAGGCGCGAGTCTTTGACCGAGGGAATCTCTGTGTAGCCTTGAAGCC 64004
|||||
QY 362 GCCAGCCCCAGAG 375
|||||
Db 64005 GCCAGCCCCAGAG 64018
|||||

RESULT 13
AC013796/c
LOCUS Homo sapiens clone RP11-20H9, LOW-PASS SEQUENCE SAMPLING.
DEFINITION AC013796
AC013796.4 GI:12583842
VERSION HTG; HTGS PHASE0.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 66980)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens, clone RP11-20H9
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 66980)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckerly,R., Boguslavskiy,L., Boukhgalter,B.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collumore,A.,
Cooke,P., Dearrellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (15-NOV-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jan 27, 2001 this sequence version replaced gi:7582586.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L3904
Center clone name: 20_H_9
-----
* NOTE: This record contains 83 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
* 1 683: contig of 683 bp in length
* 684 783: gap of 100 bp
* 784 1488: contig of 705 bp in length
* 1489 1588: gap of 100 bp
* 1589 2296: contig of 708 bp in length
```


2297 2396: gap of 100 bp
 2397 3101: contig of 705 bp in length
 3102 3201: gap of 100 bp
 3202 3916: contig of 715 bp in length
 3917 4016: gap of 100 bp
 4017 4737: contig of 721 bp in length
 4738 4837: gap of 100 bp
 4838 5557: contig of 720 bp in length
 5558 5657: gap of 100 bp
 5658 6372: contig of 715 bp in length
 6373 6472: gap of 100 bp
 6473 7155: contig of 683 bp in length
 7156 7255: gap of 100 bp
 7256 7957: contig of 702 bp in length
 7958 8057: gap of 100 bp
 8058 8773: contig of 716 bp in length
 8774 8873: gap of 100 bp
 8874 9577: contig of 704 bp in length
 9578 9677: gap of 100 bp
 9678 10392: contig of 715 bp in length
 10393 10492: gap of 100 bp
 10493 11177: contig of 687 bp in length
 11180 11279: gap of 100 bp
 11280 11984: contig of 705 bp in length
 11985 12084: gap of 100 bp
 12085 12791: contig of 707 bp in length
 12792 12891: gap of 100 bp
 12892 13600: contig of 709 bp in length
 13601 13700: gap of 100 bp
 13701 14409: contig of 709 bp in length
 14410 14509: gap of 100 bp
 14510 15221: contig of 712 bp in length
 15222 15321: gap of 100 bp
 15322 16025: contig of 704 bp in length
 16026 16125: gap of 100 bp
 16126 16822: contig of 697 bp in length
 16823 16922: gap of 100 bp
 16923 17621: contig of 699 bp in length
 17622 17721: gap of 100 bp
 17722 18423: contig of 702 bp in length
 18424 18523: gap of 100 bp
 18524 19241: contig of 718 bp in length
 19242 19341: gap of 100 bp
 19342 20056: contig of 715 bp in length
 20057 20156: gap of 100 bp
 20157 20870: contig of 714 bp in length
 20871 20970: gap of 100 bp
 20971 21671: contig of 701 bp in length
 21672 21771: gap of 100 bp
 21772 22480: contig of 709 bp in length
 22481 22580: gap of 100 bp
 22581 23293: contig of 713 bp in length
 23294 23393: gap of 100 bp
 23394 24095: contig of 702 bp in length
 24096 24195: gap of 100 bp
 24196 24916: contig of 721 bp in length
 24917 25016: gap of 100 bp
 25017 25726: contig of 710 bp in length
 25727 25826: gap of 100 bp
 25827 26534: contig of 708 bp in length
 26535 26634: gap of 100 bp
 26635 27333: contig of 699 bp in length
 27334 27433: gap of 100 bp
 27434 28131: contig of 698 bp in length
 28132 28231: gap of 100 bp
 28232 28926: contig of 695 bp in length
 28927 29026: gap of 100 bp
 29027 29758: contig of 732 bp in length
 29759 29858: gap of 100 bp
 29859 30566: contig of 708 bp in length
 30567 30666: gap of 100 bp
 30667 31368: contig of 702 bp in length
 31369 31468: gap of 100 bp

31469 32183: contig of 715 bp in length
 32184 32283: gap of 100 bp
 32284 32981: contig of 698 bp in length
 32982 33081: gap of 100 bp
 33082 33798: contig of 717 bp in length
 33799 33898: gap of 100 bp
 33899 34634: contig of 736 bp in length
 34635 34734: gap of 100 bp
 34735 35433: contig of 699 bp in length
 35434 35533: gap of 100 bp
 35534 36253: contig of 720 bp in length
 36254 36353: gap of 100 bp
 36354 37079: contig of 726 bp in length
 37080 37179: gap of 100 bp
 37180 37893: contig of 714 bp in length
 37894 37993: gap of 100 bp
 37994 38700: contig of 707 bp in length
 38701 38800: gap of 100 bp
 38801 39505: contig of 705 bp in length
 39506 39605: gap of 100 bp
 39606 40316: contig of 711 bp in length
 40317 40416: gap of 100 bp
 40417 41131: contig of 715 bp in length
 41132 41231: gap of 100 bp
 41232 41939: contig of 708 bp in length
 41940 42039: gap of 100 bp
 42040 42748: contig of 709 bp in length
 42749 42848: gap of 100 bp
 42849 43555: contig of 707 bp in length
 43556 43655: gap of 100 bp
 43656 44368: contig of 713 bp in length
 44369 44468: gap of 100 bp
 44469 45171: contig of 703 bp in length
 45172 45271: gap of 100 bp
 45272 45943: contig of 672 bp in length
 45944 46043: gap of 100 bp
 46044 46749: contig of 706 bp in length
 46750 46849: gap of 100 bp
 46850 47550: contig of 701 bp in length
 47551 47650: gap of 100 bp
 47651 48350: contig of 700 bp in length
 48351 48450: gap of 100 bp
 48451 49148: contig of 698 bp in length
 49149 49248: gap of 100 bp
 49249 49961: contig of 713 bp in length
 49962 50061: gap of 100 bp
 50062 50751: contig of 690 bp in length
 50752 50851: gap of 100 bp
 50852 51567: contig of 716 bp in length
 51568 51667: gap of 100 bp
 51668 52400: contig of 733 bp in length
 52401 52500: gap of 100 bp
 52501 53225: contig of 725 bp in length
 53226 53325: gap of 100 bp
 53326 53974: contig of 649 bp in length
 53975 54074: gap of 100 bp
 54075 54781: contig of 707 bp in length
 54782 54881: gap of 100 bp
 54882 55608: contig of 727 bp in length
 55609 55708: gap of 100 bp
 55709 56426: contig of 718 bp in length
 56427 56526: gap of 100 bp
 56527 57244: contig of 718 bp in length
 57245 57344: gap of 100 bp
 57345 58051: contig of 707 bp in length
 58052 58151: gap of 100 bp
 58152 58860: contig of 709 bp in length

Query Match 19.7% Score 74; DB 2; Length 66980;
 Best Local Similarity 100.0%; Pred. No. 3.7e-06;
 Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Db      48350 AAGAATCTCAGAGGGCGAGTCCTTGACCGAGGAATCTCTGTGTAGCTTGGAGGCC 48291
QY      362 GCCAGCCCCAGAG 375
Db      48290 GCCAGCCCCAGAG 48277

RESULT 14
LOCUS   PM12B8G/c
DEFINITION Penicillium marneffeii STS, clone pm12b8.g, sequence tagged site.
ACCESSION AL684326
VERSION   AL684326.1 GI:19337430
KEYWORDS STS.
SOURCE   Penicillium marneffeii
ORGANISM Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
          Eurotiiales; Trichocomaceae; mitosporic Trichocomaceae; Penicillium.
REFERENCE 1
AUTHORS  Yuen,K.Y., Pascal,G., Wong,S.S., Glaser,F., Woo,P.C., Kunst,F.,
          Cai,J.J., Cheung,B.Y., Medigue,C. and Danchin,A.
TITLE     Exploring the Penicillium marneffeii genome
JOURNAL   Arch. Microbiol. 179 (5), 339-353 (2003)
MEDLINE   22595073
PUBMED    12640520
REFERENCE 2 (bases 1 to 521)
AUTHORS   Danchin,A. and Pascal,G.
TITLE     Direct Submission
JOURNAL   Submitted (08-MAR-2002) Danchin A., HKU-Pasteur Research Centre,
          Dexter HC Man Building 8, Sassoon Road, Pokfulam, Hong Kong
FEATURES
    source
        1..521
            /organism="Penicillium marneffeii"
            /mol_type="genomic DNA"
            /db_xref="taxon:37727"
            /clone="pm12b8.g"

ORIGIN
Query Match      14.1%; Score 53; DB 11; Length 521;
Best Local Similarity 51.0%; Pred. No. 0.26;
Matches 122; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY      48 GGGGCGCGAGTGAAGCTGAGTTCCGGGTGGCGGGAGGCGGACTGTCCTGTGTGCTG 107
Db      317 GGGCGCGCGGGGGCGGGGGGGTGGCGGGGGCGCGCGCGCGCGCGCGCGCGCGGG 258
QY      108 AGCGCGCGGAGAGCGGCGGCGGCGGCTGATCGGCTCCCTCGAACTCGGGAGGTCCAG 167
Db      257 GCGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 198
QY      168 TGGGGTGTCTTAGGGGCCCAAGCCCGCCACCGGCTCCAAAGCTCCAGGGGCTCCCCAG 227
Db      197 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 138
QY      228 GCACCGGTGTGCGCCCTTCTTCGTGTGAGAAAGTCCGCCCGCGGGGGGAGTTCGTCCC 286
Db      137 GCGCGCGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 79

RESULT 15
LOCUS   AC148923/c
DEFINITION Alligator mississippiensis clone VMC8-211A13, WORKING DRAFT
SEQUENCE, 10 unordered pieces.
ACCESSION AC148923
VERSION   AC148923.2 GI:48427699
KEYWORDS HTG; HTGS_PHASE1; HTGS DRAFT.
SOURCE   Alligator mississippiensis (American alligator)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Archosauria; Crocodylidae; Alligatorinae; Alligator.
REFERENCE 1 (bases 1 to 134226)

```

AUTHORS

Antonellis,A., Ayele,K., Benjamin,B., Blakesley,R.W.,
 Bouffard,G.G., Brinkley,C., Brooks,S., Chu,G., Coleman,B.,
 Coleman,H., Daki,N., Engle,J., Guan,X., Gupta,J., Haghighi,P.,
 Han,J., Hansen,N., Ho,S.-L., Hu,P., Hurle,B., Idol,J.R., Jones,C.,
 Karlins,E., Kim,H., Kwong,P., Laric,P., Larson,S., Lee-Lin,S.-O.,
 Legaspi,R., Madden,M., Maduro,Q.L., Maduro,V.B., Margulies,E.H.,
 Mastaglio,C., Maskeri,B., McDowell,J., Mullikin,J.C., Faguirigan,C.,
 Park,M., Portnoy,M.E., Prasad,A., Puri,O., Reddix-Dugue,N.,
 Schandler,K., Schueier,M.G., Shah,K., Sison,C., Stantripoop,S.,
 Thomas,J.W., Thomas,P.J., Tsipouri,V., Vogt,J.L., Wetherby,K.D.,
 Young,A. and Green,E.D.

NISC Comparative Sequencing Initiative

Unpublished

2 (bases 1 to 134226)

Green,E.D.

Direct Submission

Submitted (24-APR-2004) NIH Intramural Sequencing Center, 8717
 Grovemont Circle, Gaithersburg, MD 20877, USA

3 (bases 1 to 134226)

Green,E.D.

Direct Submission

Submitted (08-JUN-2004) NIH Intramural Sequencing Center, 8717
 Grovemont Circle, Gaithersburg, MD 20877, USA

On Jun 8, 2004 this sequence version replaced gi:46559454.

COMMENT

----- Genome Center

Center: NIH Intramural Sequencing Center

Center code: NISC

Web site: <http://www.nisc.nih.gov>

Contact: nisc_zoo@nhgri.nih.gov

----- Project Information

Center project name: gng

Center clone name: 211A13

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 130266 bases at least Q40

Consensus quality: 131160 bases at least Q30

Consensus quality: 132052 bases at least Q20

Insert size: 137000; agarose-fp

Insert size: 133326; sum-of-contigs

Quality coverage: 12.67x in Q20 bases; agarose-fp

Quality coverage: 13.02x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently

* consists of 10 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1 2413: contig of 2413 bp in length

* 2414 2513: gap of unknown length

* 2514 5428: contig of 2915 bp in length

* 5429 5528: gap of unknown length

* 5529 8415: contig of 2887 bp in length

* 8416 8515: gap of unknown length

* 8516 11298: contig of 2783 bp in length

* 11299 11398: gap of unknown length

* 11399 13851: contig of 2453 bp in length

* 13852 13951: gap of unknown length

* 13952 20601: contig of 6649 bp in length

* 20601 20701: gap of unknown length

* 20701 30256: contig of 9556 bp in length

* 30257 30356: gap of unknown length

* 30357 46981: contig of 16625 bp in length

* 46982 47081: gap of unknown length

* 47082 84167: contig of 37086 bp in length

* 84168 84267: gap of unknown length

* 84268 134226: contig of 49959 bp in length.

* Location/Qualifiers

1. .134226

source

```
/organism="Alligator mississippiensis"
/mol type="genomic DNA"
/db xref="taxon:8496"
/clone="VMRC8-211A13"
/clone_lib="VMRC8"
/note="BAC resource: http://bacpac.chori.org/"
1. .2413
/note="assembly_fragment"
misc_feature
2514. .5428
/note="assembly_fragment"
misc_feature
5529. .8415
/note="assembly_fragment"
misc_feature
8516. .11298
/note="assembly_fragment"
misc_feature
11399. .13851
/note="assembly_fragment"
misc_feature
13952. .20600
/note="assembly_fragment"
misc_feature
20701. .30256
/note="assembly_fragment"
misc_feature
30357. .46981
/note="assembly_fragment"
misc_feature
47082. .84167
/note="assembly_fragment"
misc_feature
84268. .134226
/note="assembly_fragment"
clone_end:SP6
vector_side:right"
```

ORIGIN

```
Query Match 13.6%; Score 51; DB 2; Length 134226;
Best Local Similarity 51.1%; Pred. No. 0.29;
Matches 120; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

QY 13 AGCTGAGCTTACCCCTGGCGGCAACGAGCGAGGCGGGCGCGAGTGGAAAGCTGGAGTTC 72
Db 110903 AGCTGGCGGCAAGCGCGGCCCGGCGCGGCGCGGCGCGGCGGCGGCGGCGGCGG 110844

QY 73 CGGGGTGGCGGGGAGCGACTGTTCGTGTGTGTGAGCGCGCGGCGAGCGGGCGGGAG 132
Db 110843 CAGGGGGCGCGGCGAGGAGCCGACCTGTGACAGAGGGTGTGAGGGCGCCCGCGCG 110784

QY 133 CGGCTGATCGGCTCCCTCGAACTGGGGAGGTCCAGTGGGGTCCGTTAGGGCCCAAGCCC 192
Db 110783 CGCCGGCTCCGAGCCACAGGCTCGCCACAGCGCTGCATGGCGGCTGCTCTCCCGC 110724

QY 193 CCACCGGCTCCAAAGCTCCAGGGGCTCCCGCAGGACCGGTGCTCGGCCCTTC 247
Db 110723 CGCGCGCGCGCGCGCGCGCAGCATCTCCGCCGCTCGCTGCGCTGCTGC 110669
```

Search completed: April 7, 2005, 12:49:58
Job time : 1794.76 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2005, 06:13:29 ; Search time 272.403 Seconds
(without alignments)
8149.338 Million cell updates/sec

Title: US-09-977-579-4_COPY_1_375

Perfect score: 375

Sequence: 1 cctcccttcgagctgagc.....gaagccgcccagccagagaag 375

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------------|--------------------|
| 1 | 375 | 100.0 | 1261 | 3 AAC67837 | AAC67837 Human bet |
| 2 | 375 | 100.0 | 1261 | 10 ADB78651 | ADB78651 Human ion |
| 3 | 375 | 100.0 | 1261 | 10 ACF57870 | ACF57870 Human SCN |
| 4 | 369.8 | 98.6 | 4052 | 6 ABA93727 | ABA93727 Human sig |
| 5 | 288.4 | 76.9 | 953 | 13 ADS10151 | ADS10151 Human the |
| 6 | 288.4 | 76.9 | 978 | 4 AAH98320 | AAH98320 Human EST |
| 7 | 288.4 | 76.9 | 978 | 13 ADS11487 | ADS11487 Human the |
| 8 | 286.8 | 76.5 | 1045 | 4 AAK52345 | AAK52345 Human pol |
| 9 | 44.8 | 11.9 | 29340 | 13 ABD33482 | ABD33482 Human can |
| 10 | 42.6 | 11.4 | 4650 | 10 ADC32422 | ADC32422 Human nov |
| 11 | 41.2 | 11.0 | 922 | 10 ADC86708 | ADC86708 Human GPC |
| 12 | 41.2 | 11.0 | 4542 | 10 ADC30607 | ADC30607 Human nov |
| 13 | 40.2 | 10.7 | 766 | 12 ADJ39052 | ADJ39052 Plant cDN |
| 14 | 40.2 | 10.7 | 1678 | 12 ADE28355 | ADE28355 Human KPP |
| 15 | 40.2 | 10.7 | 2304 | 13 ACN43300 | ACN43300 Human dia |
| 16 | 40.2 | 10.7 | 2377 | 13 ACN43299 | ACN43299 Human dia |
| 17 | 40.2 | 10.7 | 2387 | 13 ACN43298 | ACN43298 Human dia |
| 18 | 40.2 | 10.7 | 2603 | 13 ACN43295 | ACN43295 Human dia |
| 19 | 40.2 | 10.7 | 2806 | 2 AAZ40538 | AAZ40538 Full leng |
| 20 | 40.2 | 10.7 | 2811 | 5 AAS83495 | AAS83495 DNA encod |

ALIGNMENTS

RESULT 1

AAC67837
ID AAC67837 standard; CDNA; 1261 BP.

XX AAC67837;

DT 15-FEB-2001 (first entry)

XX Human beta3 CDNA.

XX Human; beta sub-unit; beta3; analgesic; anticonvulsant;
KW cerebroprotective; vasotropic; cardiac; nootropic; cytostatic;
KW dermatological; gene therapy; voltage-gated sodium channel; pain;
KW epilepsy; stroke; ischaemia; heart disease; Jacobsen Syndrome;
KW familial nonchromaffin paraganglioma; phenylketonuria;
KW Charcot Marie tooth disease; ss.

OS Homo sapiens.

XX WO200063367-A1.

PN 26-OCT-2000.

PD 24-FEB-2000; 2000WO-EF001783.

XX 15-APR-1999; 99US-0129473P.

XX (WARN) WARNER LAMBERT CO.

PA (UYCA-) UNIV CAMBRIDGE TECH SERVICES LTD.

XX Cox P, Dixon A, Jackson A, Morgan K;

XX WPI; 2000-665241/64.

DR P-PSDB; AAB36002.

XX Novel nucleic acids encoding a beta-3 subunit from a voltage-gated sodium channel, and their corresponding polypeptides, useful for detecting and treating sodium channel-associated conditions, e.g. pain, epilepsy and stroke.

XX Claim 10; Page 70-71; 88pp; English.

XX The present sequence is given in the claims of a specification relating


```
Db      361  CGCCAGCCCCCAGAAG 375
|||||
RESULT 3
ACF57870
ID      ACF57870 standard; cDNA; 1261 BP.
XX
AC      ACF57870;
XX
DT      15-JAN-2004 (first entry)
XX
DE      Human SCN3B protein encoding cDNA.
XX
KW      SCN1A; sodium channel type 1 alpha-subunit; anticonvulsant; analgesic;
KW      neuroprotective; anesthetic; cytotatic; cerebroprotective; cardiant;
KW      hypotensive; gene therapy; SCN3B; human; gene; ss.
XX
OS      Homo sapiens.
XX
FH      Key
CD      376..1023
FT      /*tag= a
FT      /product= "SCN3B"
XX
PN      WO2003072751-A2.
XX
PD      04-SEP-2003.
XX
PF      25-FEB-2003; 2003WO-US006010.
XX
PR      25-FEB-2002; 2002US-0359382P.
XX
PA      (UVVA-) UNIV VANDERBILT.
XX
PI      George AL, Lossin C;
XX
DR      WPI; 2003-712725/67.
DR      P-PSDB; ABR83183.
XX
PT      Recombinantly expressed sodium channel type 1 alpha subunit, useful in
PT      screening for modulators, for treating e.g. epilepsy.
XX
PS      Disclosure; Page 145-147; 176pp; English.
XX
CC      The invention relates to a recombinantly expressed and isolated human
CC      SCN1A (sodium channel type 1 alpha-subunit) (I). (I), optionally
CC      incorporated into a cell, is used to screen for specific modulators,
CC      potentially useful as anticonvulsant, antiepileptic, neuroprotective,
CC      analgesic and/or anesthetic agents, e.g. for treating severe myoclonic
CC      epilepsy of infancy, stroke, cardiac arrest, hyperkalemic paralysis,
CC      motor endplate diseases, hypertension, congestive heart failure and
CC      muscular dystrophy also to treat cancer (SCN1A is expressed in prostatic
CC      and metastatic cancer cell lines). These activities can also be provided
CC      by gene therapy vectors that express (I) or the modulators. The
CC      modulators, also antibodies directed against (I), are used to detect
CC      sodium channel polypeptides. The present sequence represents a human
CC      SCN3B protein encoding cDNA
XX
SQ      Sequence 1261 BP; 263 A; 366 C; 372 G; 260 T; 0 U; 0 Other;
Query Match      100.0%; Score 375; DB 10; Length 1261;
Best Local Similarity 100.0%; Pred. No. 8e-83;
Matches 375; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1  CCCTCCCTTCGAGCTTACCTCGGGCGCAAAACGAGCGAGCGAGGGCGCGAGTGG 60
Db      1  CCCTCCCTTCGAGCTTACCTCGGGCGCAAAACGAGCGAGCGAGGGCGCGAGTGG 60
Qy      61  AAGCTGAGTTCGGGGTGGGGGAGGCGACTGTCGTTGTTCTGAGCGCCGCGAGA 120
Db      61  AAGCTGAGTTCGGGGTGGGGGAGGCGACTGTCGTTGTTCTGAGCGCCGCGAGA 120
Qy      121  GCGGGCGGAGCGGCTGATCGGCTCCCTCGAACTCGGAGGTCCAGTGGGGTGCCTTAG 180
Db      121  GCGGGCGGAGCGGCTGATCGGCTCCCTCGAACTCGGAGGTCCAGTGGGGTGCCTTAG 180
Qy      181  GGCCCAAAGCCCCCACCCGGCTCCAAAAGTCCCAAGGCTCCCGAGGCAACCGGTGCTCG 240
Db      181  GGCCCAAAGCCCCCACCCGGCTCCAAAAGTCCCAAGGCTCCCGAGGCAACCGGTGCTCG 240
Qy      241  GCCCTTCCTTCGGTCCAGAAAGTCCGCCCTGGGGGCGAGTTCGTCCTCCAAAGGTTTCCTCG 300
Db      241  GCCCTTCCTTCGGTCCAGAAAGTCCGCCCTGGGGGCGAGTTCGTCCTCCAAAGGTTTCCTCG 300
Qy      301  AAAGAATCTGAGAGGGCGCAGTCTTGACCGAGGGAATCTCTGTGTAGCTTTGGAAGC 360
Db      301  AAAGAATCTGAGAGGGCGCAGTCTTGACCGAGGGAATCTCTGTGTAGCTTTGGAAGC 360
Qy      361  CGCCAGCCCCCAGAAG 375
Db      361  CGCCAGCCCCCAGAAG 375
RESULT 4
ABA93727
ID      ABA93727 standard; cDNA; 4052 BP.
XX
AC      ABA93727;
XX
DT      30-APR-2002 (first entry)
XX
DE      Human signal transduction cDNA clone amy2_2f18.
XX
KW      Human; foetal brain; foetal kidney; melanoma; testis; amygdala;
KW      gene therapy; ss.
XX
OS      Homo sapiens.
XX
PN      WO200198454-A2.
XX
PD      27-DEC-2001.
XX
PF      25-APR-2001; 2001WO-IB002050.
XX
PR      25-APR-2000; 2000US-0199380P.
XX
PA      (GEHU-) GERMAN HUMAN GENOME PROJECT.
XX
PI      Wiemann S;
XX
DR      WPI; 2002-055860/07.
DR      P-PSDB; ABB05689.
XX
PT      Human cDNA sequences and clones derived from human fetal brain, fetal
PT      kidney, melanoma, testis and amygdala cDNA libraries, useful in genetic
PT      screening and therapy.
XX
PS      Claim 1; Page 174-175; 611pp; English.
XX
CC      The present invention describes assemblages and computer readable media
CC      comprising novel human cDNA sequences and clones derived from human
CC      foetal brain, foetal kidney, melanoma, testis and amygdala cDNA
CC      libraries. ABA93702 to ABA93766 represent human cDNA sequences from the
CC      present invention which encode the proteins given in ABB05662 to
CC      ABB05729. The human cDNA sequences and clones can be used in gene
CC      therapy. The clones may be used in a variety of applications, for example
CC      they may be used in profiling assays, for providing large arrays of human
CC      genetic material for implementing large-scale screening strategies and
CC      for treating diseases via gene therapy procedures
XX
SQ      Sequence 4052 BP; 994 A; 1033 C; 1036 G; 989 T; 0 U; 0 Other;
Query Match      98.6%; Score 369.8; DB 6; Length 4052;
Best Local Similarity 99.5%; Pred. No. 1.8e-81;
Matches 371; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

QY 3 CTCCTCTCCGAGCTGAGCTTACCTCTGGGCGCAAAACGAGCGAGCGAGGGCGCGAGTGGAA 62
 DB 431 CTCCTCTCCGAGCTGAGCTTACCTCTGGGCGCAAAACGAGCGAGCGAGGGCGCGAGTGGAA 490
 QY 63 GCTGGAGTTCCGGGCTGGGCGGAGCGACCTGCTCGTGGTCTGAGCGCGCGAGAGC 122
 DB 491 GCTGGAGTTCCGGGCTGGGCGGAGCGACCTGCTCGTGGTCTGAGCGCGCGAGAGC 550
 QY 123 GGGCGCGAGCGCTGATCGGCTCCCTCGAACTGGGGAGGTCCAGTGGGGTCTGCTTAGGG 182
 DB 551 GGGCGCGAGCGCTGATCAGCTCCCTCGAACTGGGGAGGTCCAGTGGGGTCTGCTTAGGG 610
 QY 183 CCCAAAGCCCCCACCCTGGCTCCAAAGCTCCCAAGGCTCCCGAGGCGCTCCCGAGGCGG 242
 DB 611 CCCAAAGCCCCCCTGGCTCCAAAGCTCCCAAGGCTCCCGAGGCGCTCCCGAGGCGG 670
 QY 243 CCTTCCTCTCGTTCAGAAAGTCGCCCCCTGGGGGAGTTCCTCCAAAGGTTTCCTCGAA 302
 DB 671 CCTTCCTCTCGTTCAGAAAGTCGCCCCCTGGGGGAGTTCCTCCAAAGGTTTCCTCGAA 730
 QY 303 AGAATCTGAGAGGCGCAGTCTCTTGACCGAGGGAATCTCTCTGTGTAGCTTTGGAGCGG 362
 DB 731 AGAATCTGAGAGGCGCAGTCTCTTGACCGAGGGAATCTCTCTGTGTAGCTTTGGAGCGG 790
 QY 363 CCAGCCCCAGAG 375
 DB 791 CCAGCCCCAGAG 803

RESULT 5
 AD10151
 ID ADS10151 standard; DNA; 953 BP.
 XX
 AC ADS10151;
 DT 16-DEC-2004 (first entry)
 XX
 DE Human therapeutic DNA - SEQ ID 388.
 XX
 KW antiinflammatory; neuroprotective; antianaemic; cytostatic; vulnerary;
 KW inflammatory; haematopoiesis; immunity; neurodegenerative; stem cell;
 KW aplastic anaemia; cancer; wound healing; gene therapy; ds; gene.
 XX
 OS Homo sapiens.
 XX
 PN WO2004080148-A2.
 XX
 PD 23-SEP-2004.
 XX
 PF 30-SEP-2003; 2003WO-US030720.
 XX
 PR 02-OCT-2002; 2002US-0416186P.
 XX
 PA (NUVE-) NUVELO INC.
 XX
 PI Tang YT, Asundi V, Ren F, Zhang J, Wehrman T, Wang Z, Ma Y;
 PI Wang D, Chen R, Zhao QA, Wang J, Ghosh M, Xue AJ, Weng G, Zhou P;
 XX
 WPI: 2004-668857/65.
 DR P-PSDB; ADS10835.
 XX
 PT New polynucleotide, useful in preparing a composition for diagnosing or
 PT treating inflammatory, neurodegenerative or stem cell disorders, e.g.,
 PT aplastic anemia or cancer for promoting wound healing.
 XX
 PS Claim 1; SEQ ID NO 388; 718pp; English.
 XX
 CC The invention relates to a novel isolated polynucleotide and the encoded
 CC polypeptide. The molecules of the invention demonstrate antiinflammatory,
 CC neuroprotective, antianaemic, cytostatic and vulnerary activities and may
 CC be useful in preparing a composition for diagnosing or treating
 CC inflammatory, haematopoietic, immune, neurodegenerative or stem cell

CC disorders, such as aplastic anaemia or cancer, as well as for promoting
 CC wound healing. The molecules may also be utilised during gene therapy
 CC procedures. The current sequence is that of a human therapeutic DNA of
 CC the invention. The current sequence is not shown explicitly within the
 CC specification but can be accessed from the WIPO web-site.
 XX
 SQ Sequence 953 BP; 180 A; 279 C; 290 G; 204 T; 0 U; 0 Other;
 Query Match 76.9%; Score 288.4; DB 13; Length 953;
 Best Local Similarity 99.7%; Pred. No. 1.7e-61;
 Matches 289; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 86 GAGGCGACTGTCCTGGTCTGAGCGCGCGAGAGCGGGCGGAGCGGCTGATCGGCT 145
 DB 37 GTGGCGACTGTCCTGGTCTGAGCGCGCGAGAGCGGGCGGAGCGGCTGATCGGCT 96
 QY 146 CCCTCGAACTGGGGAGGTCCAGTGGGGTCTGAGGGCCAAAGCCCCCAGCCGGTCCA 205
 DB 97 CCCTCGAACTGGGGAGGTCCAGTGGGGTCTGAGGGCCAAAGCCCCCAGCCGGTCCA 156
 QY 206 AAAGCTCCAGGCGCTCCCGAGGCGCGGTGTCGGCCCTTCTTCGGTTCAGAAAGTCGC 265
 DB 157 AAAGCTCCAGGCGCTCCCGAGGCGCGGTGTCGGCCCTTCTTCGGTTCAGAAAGTCGC 216
 QY 266 CCCCTGGGGGAGTTCGTCCTCCAAAGGTTTCCTCGAAAGAAATCTGAGAGGCGCAGTCCT 325
 DB 217 CCCCTGGGGGAGTTCGTCCTCCAAAGGTTTCCTCGAAAGAAATCTGAGAGGCGCAGTCCT 276
 QY 326 TGACCGAGGGAATCTCTGTGTAGCTTTGGAAGCGCCAGCCGCCAGAG 375
 DB 277 TGACCGAGGGAATCTCTGTGTAGCTTTGGAAGCGCCAGCCGCCAGAG 326

RESULT 6
 AAH98320
 ID AAH98320 standard; cDNA; 978 BP.
 XX
 AC AAH98320;
 DT 12-OCT-2001 (first entry)
 XX
 DE Human EST-derived coding sequence SEQ ID NO: 177.
 XX
 KW Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
 KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
 KW diagnostics; forensic test; gene mapping; genetic disorder; biodiversity;
 KW gene therapy; nutrition; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200154477-A2.
 XX
 PD 02-AUG-2001.
 XX
 PF 25-JAN-2001; 2001WO-US002687.
 XX
 PR 25-JAN-2000; 2000US-00491404.
 PR 17-JUL-2000; 2000US-00617746.
 PR 03-AUG-2000; 2000US-00631451.
 PR 15-SEP-2000; 2000US-00663870.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
 PI Cao Y, Drmanac RA, Zhang J, Werhman T;
 XX
 WPI: 2001-476164/51.
 DR P-PSDB; AAM23661.
 XX
 PT Isolated polypeptide for treatment of diseases, diagnostics, raising
 PT antibodies and research use.
 XX
 PS Claim 1; Page 299-300; 1275pp; English.

XX The present invention provides the protein and coding sequences of novel
CC proteins from a variety of organisms, including human, dog, cat, horse,
CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
CC urchin and tomato. These were derived from expressed sequence tags (ESTs)
CC from the organism of interest. They can be used in diagnostics,
CC forensics, gene mapping, identification of mutations, to assess
CC biodiversity and for nutritional purposes. The present sequence is a cDNA
CC of the invention
XX
SQ Sequence 978 BP; 193 A; 283 C; 293 G; 209 T; 0 U; 0 Other;

Query Match 76.9%; Score 288.4; DB 4; Length 978;
Best Local Similarity 99.7%; Pred. No. 1.7e-61;
Matches 289; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 86 GAGCGCACTGTCGTGCTGAGCGCGGCGAGCGGCGGCGGCGGCGGCTGATCGGCT 145
DB 5 GTGCGCACTGTCGTGCTGAGCGCGGCGGCGGCGGCGGCGGCGGCTGATCGGCT 64
QY 146 CCCTCGAACTGGGAGGTCAGTGGGTCGCTTAGGGCCCAAGCCCCCGGCTCCA 205
DB 65 CCCTCGAACTGGGAGGTCAGTGGGTCGCTTAGGGCCCAAGCCCCCGGCTCCA 124
QY 206 AAAGCTCCAGGGCTCCCGAGGACCGGTGCTCGGCGCTTCTTCGCTCAGAAAGTCGC 265
DB 125 AAAGCTCCAGGGCTCCCGAGGACCGGTGCTCGGCGCTTCTTCGCTCAGAAAGTCGC 184
QY 326 TGACCGAGGAATCTCTCTGTAGCCTTGAAGCCCGCCAGCCAGAG 375
DB 245 TGACCGAGGAATCTCTCTGTAGCCTTGAAGCCCGCCAGCCAGAG 294

RESULT 7
ADSL11487
ID ADSL11487 standard; DNA; 978 BP.
XX AC ADSL11487;
XX
DT 16-DEC-2004 (first entry)
XX
DE Human therapeutic contig DNA - SEQ ID 1724.
XX
KW antiinflammatory; neuroprotective; antianaemic; cytostatic; vulnery;
KW inflammatory; haematopoiesis; immunity; neurodegenerative; stem cell;
KW aplastic anaemia; cancer; wound healing; gene therapy; ds; gene.
XX
OS Homo sapiens.
XX
FN WO2004080148-A2.
XX
PD 23-SEP-2004.
XX
PF 30-SEP-2003; 2003WO-US030720.
XX
PR 02-OCT-2002; 2002US-0416186P.
XX
PA (NUVE-) NUVELO INC.
XX
PI Tang YT, Asundi V, Ren F, Zhang J, Zhang J, Wehrman T, Wang Z, Ma Y;
PI Wang D, Chen R, Zhao QA, Wang J, Ghosh M, Xue AU, Weng G, Zhou P;
XX WPI; 2004-668857/65.
DR P-PSDB; ADS12085.
XX
PT New polynucleotide, useful in preparing a composition for diagnosing or
PT treating inflammatory, neurodegenerative or stem cell disorders, e.g.,
PT aplastic anemia or cancer for promoting wound healing.
XX

PS Example 2; SEQ ID NO 1724; 718pp; English.
XX
CC The invention relates to a novel isolated polynucleotide and the encoded
CC polypeptide. The molecules of the invention demonstrate antiinflammatory,
CC neuroprotective, antianaemic, cytostatic and vulnerary activities and may
CC be useful in preparing a composition for diagnosing or treating
CC inflammatory, haematopoietic, immune, neurodegenerative or stem cell
CC disorders, such as aplastic anaemia or cancer, as well as for promoting
CC wound healing. The molecules may also be utilised during gene therapy
CC procedures. The current sequence is that of a human therapeutic contig
CC DNA of the invention. The current sequence is not shown explicitly within
CC the specification but can be accessed from the WIPO web-site.
XX
SQ Sequence 978 BP; 193 A; 283 C; 293 G; 209 T; 0 U; 0 Other;

Query Match 76.9%; Score 288.4; DB 13; Length 978;
Best Local Similarity 99.7%; Pred. No. 1.7e-61;
Matches 289; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 86 GAGCGCACTGTCGTGCTGAGCGCGGCGAGCGGCGGCGGCGGCTGATCGGCT 145
DB 5 GTGCGCACTGTCGTGCTGAGCGCGGCGGCGGCGGCGGCGGCTGATCGGCT 64
QY 146 CCCTCGAACTGGGAGGTCAGTGGGTCGCTTAGGGCCCAAGCCCCCGGCTCCA 205
DB 65 CCCTCGAACTGGGAGGTCAGTGGGTCGCTTAGGGCCCAAGCCCCCGGCTCCA 124
QY 206 AAAGCTCCAGGGCTCCCGAGGACCGGTGCTCGGCGCTTCTTCGCTCAGAAAGTCGC 265
DB 125 AAAGCTCCAGGGCTCCCGAGGACCGGTGCTCGGCGCTTCTTCGCTCAGAAAGTCGC 184
QY 266 CCCTGGGGGCGAGTTCGTCCTCGAAAGGTTTCCTCGAAAGAAATCTGAGAGGCGGCGAGTCCT 325
DB 185 CCCTGGGGGCGAGTTCGTCCTCGAAAGGTTTCCTCGAAAGAAATCTGAGAGGCGGCGAGTCCT 244
QY 326 TGACCGAGGAATCTCTCTGTAGCCTTGAAGCCCGCCAGCCAGAG 375
DB 245 TGACCGAGGAATCTCTCTGTAGCCTTGAAGCCCGCCAGCCAGAG 294

RESULT 8
AAK52345
ID AAK52345 standard; cDNA; 1045 BP.
XX AC AAK52345;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 890.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
FN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US004098.
XX
PR 03-FEB-2000; 2000US-00496914.
PR 27-APR-2000; 2000US-00560875.
PR 20-JUN-2000; 2000US-00598075.
PR 19-JUL-2000; 2000US-00620325.
PR 01-SEP-2000; 2000US-00654936.
PR 15-SEP-2000; 2000US-00663561.
PR 20-OCT-2000; 2000US-00693325.
PR 30-NOV-2000; 2000US-00728422.
XX
PA (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
 PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 XX WPI: 2001-476283/51.
 DR P-PSDB; AAM79212.
 XX Nucleic acids encoding polypeptides with cytokine-like activities, useful
 in diagnosis and gene therapy.
 XX Claim 1; Page 2934-2935; 6221pp; English.
 XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
 encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
 cytokine, cell proliferation or cell differentiation or which may induce
 production of other cytokines in other cell populations. The
 polynucleotides and polypeptides are useful in gene therapy, vaccines or
 peptide therapy. The polypeptides have various cytokine-like activities,
 e.g. stem cell growth factor activity, haematopoiesis regulating
 activity, tissue growth factor activity, immunomodulatory activity and
 activin/inhibin activity and may be useful in the diagnosis and/or
 treatment of cancer, leukaemia, nervous system disorders, arthritis and
 inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
 (AAK52582) and 3656 (AAM80020) are omitted as the relevant pages from the
 sequence listing were missing at the time of publication
 XX Sequence 1045 BP; 222 A; 289 C; 314 G; 220 T; 0 U; 0 Other;
 Query Match 76.5%; Score 286.8; DB 4; Length 1045;
 Best Local Similarity 99.3%; Pred. No. 4.2e-51;
 Matches 288; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 86 GAGGCGACTGTCCGTGCTGCTGAGCCCGCGGAGAGCGGGCGGAGCGGCTGATCGGCT 145
 Db 12 GTGGCGACTGTCCGTGCTGCTGAGCCCGCGGAGAGCTGGCGGAGCGGCTGATCGGCT 71
 QY 146 CCCTCGAATCGGGAGGTTCAGTGGGTGCTGTAGGGCCCAAGCCCCACCGGCTCA.205
 Db 72 CCCTCGAATCGGGAGGTTCAGTGGGTGCTGTAGGGCCCAAGCCCCACCGGCTCA 131
 QY 206 AAAGCTCCAGGGCCCTCCCGAGSCCGGTGCTCGGCCCTTCTTCGTCAGAAAGTCGC 265
 Db 132 AAAGCTCCAGGGCCCTCCCGAGSCCGGTGCTCGGCCCTTCTTCGTCAGAAAGTCGC 191
 QY 266 CCCTCGGGGCGAGTTCGTCCTCCAAAGGGTTTCCTCGAAAGAAATCTGAGAGGGCGCAGTCT 325
 Db 192 CCCTCGGGGCGAGTTCGTCCTCCAAAGGGTTTCCTCGAAAGAAATCTGAGAGGGCGCAGTCT 251
 QY 326 TGACCGAGGGAATCTCTGTGTAGCTTGGAAAGCCGCCAGCCCAAGAAG 375
 Db 252 TGACCGAGGGAATCTCTGTGTAGCTTGGAAAGCCGCCAGCCCAAGAAG 301
 RESULT 9
 ABD33482
 ID ABD33482 standard; DNA; 29340 BP.
 XX AC
 AC ABD33482;
 XX DT
 DT 18-NOV-2004 (first entry)
 XX DE
 DE Human cancer-associated (CA) gene HD07-093.
 XX Human; cancer-associated protein; CAP; cancer-associated gene; CA; gene;
 KW Human; cancer-associated protein; CAP; cancer-associated gene; CA; gene;
 KW da; cancer; cytostatic.
 XX KW
 KW Homo sapiens.
 OS
 OS WO2004058146-A2.
 PN
 PN 15-JUL-2004.

PF 15-DEC-2003; 2003WO-US040081.
 XX 17-DEC-2002; 2002US-00322281.
 PR (SAGR-) SAGRES DISCOVERY INC.
 XX PA
 PA Morris DW, Malandro MS;
 XX PI
 PI WPI; 2004-499109/47.
 DR Novel human cancer associated protein encoded within open reading frame
 of cancer associated gene, useful as targets for diagnosing cancer.
 XX PT
 PT Claim 16; SEQ ID NO 642; 182pp; English.
 XX The invention relates to cancer-associated proteins (CAP) and the cancer-
 associated (CA) nucleic acids encoding them. The invention also relates
 to a method for treating cancers involving administering to a patient an
 inhibitor of CAP, and a method of screening for anticancer activity in a
 potential drug involving providing a cell that expresses a CA gene,
 contacting a tissue sample derived from a cancer cell with an anticancer
 drug candidate and monitoring the effect of the anticancer drug candidate
 on expression of the CA gene. The CAP proteins are useful for detecting
 cancer associated with expression of a CAP protein in a test cell sample
 and for screening for a bioactive agent capable of modulating the
 activity of a CAP protein. The CA nucleic acids are useful for diagnosing
 cancer, involving determining the expression of a CA nucleic acid in a
 tissue. This sequence represents a human CA gene of the invention. Note:
 The sequence data for this patent did not form part of the printed
 specification, but was obtained in electronic format directly from WIPO
 at ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 29340 BP; 6351 A; 8588 C; 8017 G; 5384 T; 0 U; 0 Other;
 Query Match 11.9%; Score 44.8; DB 13; Length 29340;
 Best Local Similarity 56.7%; Pred. No. 0.28;
 Matches 102; Conservative 0; Mismatches 77; Indels 1; Gaps 1;
 QY 74 GGGGTGGGGGGAGGCGACTGTCGTGTGTGAGCCCGGAGAGCGGGCGGAGC 133
 Db 9834 GGGAGGGGGGGTAGGGGACCTTTTCAGAGCCAGGAGGGGTTTCGGGGGGCGTGGGGGGCGC 9893
 QY 134 GCCTGATCGGCTCCCTCGAATCGGGGAGGTCCAGTGGGGTCG-CTTAGGGCCCAAGCC 192
 Db 9894 TCGGAGCGGAGCCCGGGCTCGACGGCGTGCCTGGCGCGAGTGTATGACAGCGCGC 9953
 QY 193 CCACCGGGCTCAAAAGTCCAGGGCTCCCGAGGACCCGAGCGTGTGCGCCCTTCCTTCG 252
 Db 9954 CCGGCCCGAAACCCCGAGCCCGGGGCTCCCGACCGCGGCGCTCCCGCCCTCCCGCG 10013
 RESULT 10
 ADC32422
 ID ADC32422 standard; cDNA; 4650 BP.
 XX AC
 AC ADC32422;
 XX DT
 DT 18-DEC-2003 (first entry)
 XX DE
 DE Human novel cDNA contig sequence, SEQ ID NO:2504.
 XX Human; diagnostic; drug screening; forensics; gene mapping;
 KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
 KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
 KW ulcers; osteoporosis; autoimmune disease; cancer;
 KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
 KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;
 KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
 KW gene therapy; chromosome 20; ss.
 XX OS
 OS Homo sapiens.
 XX PN
 PN WO2003029271-A2.

XX PD 10-APR-2003.
XX PF 24-SEP-2002; 2002WO-US030474.
XX PR 24-SEP-2001; 2001US-0324631P.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
XX PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
XX PI Haley-Vicente D, Drmanac RT;
XX DR WPI; 2003-371981/35.
XX DR P-PSDB; ADC33189.
XX PT New polynucleotide and polypeptide useful for diagnosing, preventing or
XX PT treating conditions such as neurodegenerative diseases, anemias, platelet
XX PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
XX PT cancer.
XX PS Example 2; SEQ ID NO 2504; 1185pp; English.
XX CC The invention relates to 971 novel human cDNA sequences (ADC29919-
CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
CC invention also relates to nucleic acid sequences over 99% identical with
CC the novel human cDNAs. The invention additionally encompasses expression
CC vectors and host cells comprising a nucleic acid of the invention; the
CC recombinant production of a polypeptide of the invention; an antibody
CC against a polypeptide of the invention; a method of detecting
CC polynucleotides or polypeptides of the invention; and methods of
CC identifying a compound which binds to a polypeptide of the invention. The
CC invention further discloses methods of preventing, treating or
CC ameliorating a medical condition; kits comprising polynucleotide probes
CC and/or monoclonal antibodies for carrying out the methods of the
CC invention; methods for the identification of compounds that modulate the
CC expression or activity of the polynucleotide and/or polypeptide; and 767
CC contig sequences corresponding to the cDNA sequences of the invention
CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
CC -ADC33394). The nucleic acids and polypeptides of the invention are
CC useful in diagnostics, drug screening, forensics, gene mapping, in the
CC identification of mutations responsible for genetic disorders or other
CC traits, for assessing biodiversity, and in producing many other types of
CC data and products dependent on DNA and amino acid sequences. They are
CC also used for treating diseases such as Parkinson's disease, Alzheimer's
CC disease and other neurodegenerative diseases, anaemia, platelet
CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
CC cancer. The nucleic acids may also be used as hybridisation probes or
CC primers, and in the recombinant production of a protein. The polypeptides
CC are also useful in generating antibodies, as molecular weight markers,
CC and as food supplements. The present sequence represents a human contig
CC sequence used in an example of the invention. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 4650 BP; 911 A; 1214 C; 1469 G; 1056 T; 0 U; 0 Other;
Query Match 11.4%; Score 42.6; DB 10; Length 4650;
Best Local Similarity 54.9%; Pred. No. 0.77;
Matches 84; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
QY 10 CCAGAGCTGAGCTTACCTCGGCGCAACAGCGAGCGAGCGGGCGCGAGCTGGAG 69
Db 3729 CGGCGCTGCCCCCGCTTCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGGGG 3788
QY 70 TTCGGGGGTGGCGGGGAGCGAGCTGTCCGTGGTGTGTGAGCGCGCGCGAGCGGGCGCG 129
Db 3789 GTCCGGCGGGGTGGCGCGGGCGCGCGCGCGCGAGAGTGGCGGGCGCGGGGGGCG 3848
QY 130 GAGCGGCTGATCGGCTCCCTCGAAGCTGGGAGG 162
Db 3849 GGGCGGCGGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3881

RESULT 11
ADC86708
ID ADC86708 standard; DNA; 922 BP.
XX AC ADC86708;
XX DT 01-JAN-2004 (first entry)
XX DE Human GPCR gene SEQ ID NO:1161.
XX KW ds; gene; human; GPCR;
XX KW guanosine triphosphate-binding protein coupled receptor; gene therapy.
XX OS Homo sapiens.
XX EP EP1270724-A2.
XX PD 02-JAN-2003.
XX PF 18-JUN-2002; 2002EP-00013517.
XX PR 18-JUN-2001; 2001JP-00246789.
XX PA (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
XX PA (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
XX PI Suwa M, Asai K, Akiyama Y, Aburatani H;
XX WPI; 2003-315783/31.
XX DR P-PSDB; ADC86709.
XX PT New polynucleotide, useful for preparing a composition for treating a
XX PT patient in need of increased or suppressed activity or expression of the
XX PT guanosine triphosphate-binding protein coupled receptor.
XX PS Claim 1; SEQ ID NO 1161; 28pp; English.
XX CC The invention relates to a novel polynucleotide encoding a guanosine
XX CC triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of
XX CC the invention may have a use in gene therapy. The polynucleotide and
XX CC polypeptide are useful for preparing a composition for treating a patient
XX CC in need of increased or suppressed activity or expression of the
XX CC guanosine triphosphate-binding protein coupled receptor. The
XX CC polynucleotide sequences shown in ADC85548-ADC87616 encode GPCR's of the
XX CC invention.
XX SQ Sequence 922 BP; 118 A; 82 C; 356 G; 106 T; 0 U; 260 Other;
Query Match 11.0%; Score 41.2; DB 10; Length 922;
Best Local Similarity 40.9%; Pred. No. 1.4;
Matches 70; Conservative 0; Mismatches 101; Indels 0; Gaps 0;
QY 12 GAGCTGAGCTTACCTCGGCGCAACAGCGAGCGAGCGGGCGCGAGTGGAGCTGGAGTT 71
Db 363 GCGNN 422
QY 72 CCGGGGTGGCGGGGAGGCGACTGTCCGTGTGTGTGAGCGCCCGCGAGAGCGGGCGCGGA 131
Db 423 GGGGGGGGGGGGGGNNNGNNGNNGNNGNNGNNGNNGNNGNNGNNGNNGNNGNNGG 482
QY 132 GCGGCTGATCGGCTCCCTCGAACTGGGAGGTCCAGTGGGGTCCGTTAGG 182
Db 483 GGGGNN 533
RESULT 12
ADC30607/c
ID ADC30607 standard; cDNA; 4542 BP.
XX AC ADC30607;
XX XX

DT 18-DEC-2003 (first entry)
XX Human novel cDNA sequence, SEQ ID NO:689.
DE
XX
XX Human; diagnostic; drug screening; forensics; gene mapping;
KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KW neurodegenerative diseases; anemia; platelet disorder; wound; burns;
KW ulcers; osteoporosis; autoimmune disease; cancer;
KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
KW neuroprotective; antianemic; anticoagulant; thrombolytic; vulnerary;
KW antitumor; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
KW gene therapy; chromosome 20; gene; ss.
XX
XX Homo sapiens.
OS
XX WO2003029271-A2.
PN
XX
XX 10-APR-2003.
PD
XX 24-SEP-2002; 2002WO-US030474.
PF
XX 24-SEP-2001; 2001US-0324631P.
PR
XX (HYSE-) HYSEQ INC.
PA
XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
PI Haley-Vicente D, Drmanac RT;
XX
XX WPI: 2003-371981/35.
DR P-PSDB; ADC31578.
DR
XX New polynucleotide and polypeptide useful for diagnosing, preventing or
PT treating conditions such as neurodegenerative diseases, anemias, platelet
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
PT cancer.
XX
XX Claim 1; SEQ ID NO 689; 1185pp; English.
PS
XX The invention relates to 971 novel human cDNA sequences (ADC29919-
CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
CC invention also relates to nucleic acid sequences over 99% identical with
CC the novel human cDNAs. The invention additionally encompasses expression
CC vectors and host cells comprising a nucleic acid of the invention; the
CC recombinant production of a polypeptide of the invention; an antibody
CC against a polypeptide of the invention; a method of detecting
CC polynucleotides or polypeptides of the invention; and methods of
CC identifying a compound which binds to a polypeptide of the invention. The
CC invention further discloses methods of preventing, treating or
CC ameliorating a medical condition; kits comprising polynucleotide probes
CC and/or monoclonal antibodies for carrying out the methods of the
CC invention; methods for the identification of compounds that modulate the
CC expression or activity of the polynucleotide and/or polypeptide; and 767
CC contig sequences corresponding to the cDNA sequences of the invention
CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
CC -ADC33394). The nucleic acids and polypeptides of the invention are
CC useful in diagnostics, drug screening, forensics, gene mapping, in the
CC identification of mutations responsible for genetic disorders or other
CC traits, for assessing biodiversity, and in producing many other types of
CC data and products dependent on DNA and amino acid sequences. They are
CC also used for treating diseases such as Parkinson's disease, Alzheimer's
CC disease and other neurodegenerative diseases, anemia, platelet
CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
CC cancer. The nucleic acids may also be used as hybridisation probes or
CC primers, and in the recombinant production of a protein. The polypeptides
CC are also useful in generating antibodies, as molecular weight markers,
CC and as food supplements. The present sequence represents a specifically
CC claimed human cDNA sequence of the invention. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 4542 BP; 1049 A; 1409 C; 1182 G; 902 T; 0 U; 0 Other;
SQ

Query Match 11.0%; Score 41.2; DB 10; Length 4542;
Best Local Similarity 56.7%; Pred. No. 1.7;
Matches 76; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
QY 10 CCGAGCTGAGCTTACCTGGGCGCAACGAGCGAGCGAGGGCGGAGTGGAGCTGGAG 69
DB 826 CGGCGCTGCCCCCTTCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 767
QY 70 TTCCCGGGGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 129
DB 766 GTCCGCGGGGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 707
QY 130 GAGCGGCTGATCGG 143
DB 706 GGGCGGCGGCGGCGG 693
RESULT 13
ADJ39052/c
ID ADJ39052 standard; cDNA; 766 BP.
XX
AC ADJ39052;
XX
DT 06-MAY-2004 (first entry)
XX
DE Plant cDNA #52.
XX
KW Plant; gene; ss; transcription; plant genome augmentation; cereal;
KW soybean; alfalfa; sunflower; canola; cotton; peanut; tobacco; sugar beet;
KW maize; barley; sorghum; rice; wheat; crop plant; insecticide resistance;
KW stress tolerance; salt tolerance; cold tolerance; drought tolerance;
KW plant nutrition; apical dominance; dwarfism; early flowering; antiviral;
KW antifungal.
XX
OS Eukaryota.
XX
PN US2004016025-A1.
XX
PD 22-JAN-2004.
XX
PF 26-SEP-2002; 2002US-00260238.
XX
PR 26-SEP-2001; 2001US-0325277P.
PR 26-SEP-2001; 2001US-0325448P.
PR 04-APR-2002; 2002US-0370620P.
XX
XX (BUDW/) BUDWORTH P.
PA (MOUG/) MOUGHAMER T.
PA (BRIG/) BRIGGS S P.
PA (COOP/) COOPER B.
PA (GLAZ/) GLAZEBROOK J.
PA (GOLF/) GOLF S A.
PA (KATA/) KATAGIRI F.
PA (KREP/) KREPS J.
PA (PROV/) PROVANT N.
PA (RICK/) RICK D.
PA (ZHUT/) ZHU T.
XX
XX Budworth P, Moughamer T, Briggs SP, Provant N, Ricke D, Zhu T;
PI Goff SA, Katagiri F, Kreps J, Moughamer T, Ricke D, Zhu T;
XX WPI: 2004-190374/18.
DR
XX New rice promoter, useful for manipulating crop plants to alter or
PT improve phenotypic characteristics, e.g. produce large quantities of oil
PT or proteins, resistance to insecticides, virus or fungi, stress tolerance
PT or high nutritional value.
XX
XX Claim 68; SEQ ID NO 52; 230pp; English.
PS
XX The invention relates to plant nucleotide sequences that direct seed-,
CC leaf- and/or stem-, panicle-, root- or pollen-specific or -preferential


or constitutive transcription of an operatively linked nucleic acid segment. The invention also relates to a method for augmenting a plant genome and a method of identifying a gene, where its expression is altered in the seed, leaf, stem, panicle, pollen, root or is constitutive in a plant cell. The plant is a cereal, e.g. soybean, alfalfa, sunflower, canola, cotton, peanut, tobacco or sugar beet, preferably maize, barley, sorghum, rice or wheat. The polynucleotides and the polypeptides they encode are useful for manipulating crop plants to alter or improve phenotypic characteristics, to produce large quantities of oil or proteins, to incur resistance to insecticides, viruses or fungi, and to incur stress tolerance (e.g. salt, cold or drought) to ensure the plants have a high nutritional value with reduced apical dominance or dwarfism, early flowering or altered metabolic pathways. This sequence represents a plant nucleic acid of the invention. Note: The sequence data for this patent did not form part of the printed specification but was obtained in patent data directly from USPTO at seqdata.uspto.gov/sequence.html.

SQ Sequence 766 BP; 89 A; 343 C; 192 G; 141 T; 0 U; 1 Other;

| | | | | |
|-----------------------|--------|----------------|--------|----------------|
| Query Match | 10.7% | Score 40.2; | DB 12; | Length 766; |
| Best Local Similarity | 57.6%; | Pred. No. 2.4; | | |
| Matches | 72; | Conservative | 0; | Mismatches 53; |
| | | | | Indels 0; |
| | | | | Gaps 0; |

| | | | |
|----|-----|--|-----|
| Qy | 12 | GAGCTGAGTTACCTTGGGCGCAACGACGAGGAGCGGGGCGCGAGTGGAAAGTGGAGTT | 71 |
| | | | |
| Db | 686 | GAGCGCGCGCGCGTTCGCGGAGCGAGGTGAGGCGAGCGGCGGACGAGTGGCGCGGT | 62 |
| | | | |
| Qy | 72 | CGGGGTGGGCGGGGAGGCGCACTGTCCGTGTGCTGAGCGCGCGGAGAGCGGCGCGGA | 13 |
| | | | |
| Db | 626 | CGGTTCGCGGACGAGGCGGCGCACTGTGTCGTGGAGGCGAGGAGAGCGGCGGCGCGGA | 56 |
| | | | |
| Qy | 132 | CGGGC | 136 |
| | | | |
| Db | 566 | TGAGC | 562 |

RESULT 14

| ADE28355 | |
|---|-----------------------------------|
| ID | ADE28355 standard: cDNA: 1678 BP. |
|  | |

AC ADE28355;

29-JAN-2004 (first entry)

AA
DE Human KPP cDNA - SEO ID 66.

kinase; phosphatase; KPP; hepatotropic; antiarteriosclerotic;
antipsoriatic; cytosstatic; haemostatic; muscular; cerebroprotective;
neotropic; ophthalmological; anticonvulsant; vasotropic; neuroprotective;
antiparkinsonian; antiasthmatic; antianaemic; antiasthmatic;
antidiabetic; antiinflammatory; osteopathic; antiautismic;
antirheumatic; dermatocidal; virucide; antibacterial; fungicide;
antiparasitic; protozoacide; antihelminthic; antitoux; cardiovascular;
antiarteriosclerotic; immunosuppressive; cell proliferative; cirrhosis;
hepatitis; arteriosclerosis; psoriasis; primary thrombocytopenia; cancer;
developmental; renal tubular acidosis; Becker's muscular dystrophy;
gonadal dysgenesis; hypothyroidism; seizure; neurological;
Fick's disease; cataract; epilepsy; ischaemic cerebrovascular; stroke;
Alzheimer's; Parkinson's; dementia; autoimmune; inflammatory; AIDS;
allergy; anaemia; asthma; diabetes mellitus; bronchitis; osteoporosis;
osteoarthritis; rheumatoid arthritis; contact dermatitis; gout;
lipid disorder; cholestasis; Gaucher's; diabetes; atherosclerosis; liver;
viral; bacterial; fungal; parasitic; protozoan; helminthic infection;
trauma; gene therapy; human; ss; gene.

OS Homo sapiens.

AA
PN
WO2003080805-A2

02-OCT-2003

18-MAR-2003: 2003WO-US008715.

PR 19-MAR-2002; 2002US-0366088P.
PR 29-MAR-2002; 2002US-0369248P.

PR 29-MAR-2002; 2002US-0369248P.

PA (INCY-) INCYTE CORP.

Chien D, Jin P, Hawkins PR, Baughn MR, Becha SD, Chang H;
 PI Ding L, Elliott VS, Emerling BM, Gandhi AR, Gietzen KJ, Griffin JA;
 PI Gururajan R, Hafalia AJA, Ison CH, Kable AE, Khare R, Lee SY;
 PI Lee EA, Lu Y, Marquis JP, Lehr-Nason PM, Runkumar J, Richardson TW;
 PI Swarnakar A, Tran UK, Chawla NK, Yao MG, Yue H, Bhatia U;
 PI Burrill JD, Lee S, Blake JW, Ho A, Zheng W;

DR WPI; 2004-011523/01.

DR P-PSDB; ADE28303.

New human kinases and phosphatases, and polynucleotides encoding them useful for treating, preventing or diagnosing e.g. cell proliferative disorders, inflammatory, autoimmune, viral, bacterial, parasitic or fungal diseases.

PS Claim 5: SEQ ID NO 66: 340pp: English.

The invention relates to a novel isolated kinase and phosphatase (KPP) polypeptide. The polypeptide of the invention demonstrates hepatotropic, antiarteriosclerotic, antipsoriatic, cytostatic, haemostatic, muscular, vasoprotective, nootropic, ophthalmological, anticonvulsant, caesotropic, neuroprotective, antiparkinsonian, antiasthmatic, antianemic, antiasthmatic, antidiabetic, antiinflammatory, osteopathic, antiarthritic, antirheumatic, dermatological, virucide, antibacterial, fungicide, antiparasitic, protozoacide, antihelminthic, antigout, cardiovascular, antiarteriosclerotic and immunosuppressive activities. The KPP polypeptides may be useful for diagnosing, treating or preventing cell proliferative disorders including cirrhosis, hepatitis, arteriosclerosis, psoriasis, primary thrombocytopaenia and cancer. developmental disorders such as renal tubular acidosis, Becker's muscular dystrophy, gonadal dysgenesis, hypothyroidism or seizures, neurological disorders e.g. Pick's disease, cataract, epilepsy, ischaemic cerebrovascular disease, stroke, Alzheimer's disease, Parkinson's disease or dementia, autoimmune or inflammatory disorders including AIDS, allergies, anaemia, asthma, diabetes mellitus, bronchitis, osteoporosis, osteoarthritis, rheumatoid arthritis, contact dermatitis or gout and lipid disorders such as cholestasis, Gaucher's disease, diabetes, arteriosclerosis or liver disease, as well as viral, bacterial, fungal, parasitic, protozoan or helminthic infections and trauma. Furthermore, the polypeptide may be utilised during gene therapy procedures. The current sequence is that of the human KPP cDNA of the invention.

Sequence 1678 BP; 316 A; 553 C; 527 G; 282 T; 0 U; 0 Other;

| | | | | |
|---------------------------|-------|-----------------|-----------|--------------|
| Query Match | 10.7% | Score 40.2; | DB 12; | Length 1678; |
| Best Local Similarity | 49.8% | Pred. No. 2.7; | | |
| Matches 102; Conservative | 0; | Mismatches 103; | Indels 0; | Gaps 0 |

| | | | |
|----|-----|---|----|
| Qy | 23 | ACCTCGGGCGCAACGACGAGCAGGGCGCGAGTGTGAACTGTGAGTTTCGGGGTGGGC | 82 |
| Db | 593 | ACACCGCCAGAGGGGGCCACGGGAAGCAGCAGCCGAGGCCGTTCGCCGTTCACAGC | 65 |
| Qy | 83 | GGGAGCGCACTGTTCGTGGTCTGACGCCCGCGCAGACGGCGCGAGCGGTGTGTCG | 14 |
| Db | 653 | GAGCGGGTGGCGGCAGTGTGTACAGGCGACGGCGGGGCCAGAGAAAGAGGCCCAAGTCT | 71 |
| Qy | 143 | GCTCCCTCGAACTGGGGGAGGTCCAGTGGGGTGCCTTAGGGCCCAAAGCCCCACC | 20 |
| Db | 713 | TCCAGGGAGGGCTCAGGGGGTCCACAGAGTCTCTCCGGGACAAACGCCCTCTCCGGG | 77 |

QV : 203 CCAAAGCTCCACGGGCTCCCCAG 227

773 CCTGATGTGGCACCCCCAGCCTG 797

RESULT 15

ACN4330

ACN43300
ID ACN43300 standard; cDNA; 2304 BP.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2005, 08:55:04 ; Search time 86.5385 Seconds
(without alignments)
7090.532 Million cell updates/sec

Title: US-09-977-579-4_COPY_1_375

Perfect score: 375

Sequence: 1 cctcccttcgagcgagc.....gaagccgagccagccagagaag 375

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|----------------------|
| 1 | 44.8 | 11.9 | 11558 | 5 | PCT-US93-06251-23 |
| 2 | 42 | 11.2 | 9622 | 4 | US-09-949-016-15003 |
| 3 | 40.4 | 10.8 | 6651 | 4 | US-09-902-540-4944 |
| 4 | 40.4 | 10.8 | 30780 | 4 | US-09-902-540-1243 |
| 5 | 40.2 | 10.7 | 2802 | 4 | US-09-949-016-1335 |
| 6 | 40.2 | 10.7 | 2806 | 4 | US-09-688-188B-102 |
| 7 | 40.2 | 10.7 | 2806 | 4 | US-09-231-117D-102 |
| 8 | 40.2 | 10.7 | 2838 | 4 | US-09-949-016-794 |
| 9 | 40.2 | 10.7 | 57559 | 4 | US-09-949-016-13077 |
| 10 | 40.2 | 10.7 | 57560 | 4 | US-09-949-016-12536 |
| 11 | 40 | 10.7 | 21295 | 4 | US-09-902-540-1194 |
| 12 | 39.2 | 10.5 | 601 | 4 | US-09-949-016-124631 |
| 13 | 39.2 | 10.5 | 8374 | 4 | US-09-949-016-15257 |
| 14 | 39.2 | 10.5 | 45225 | 4 | US-09-949-016-12428 |
| 15 | 39.2 | 10.5 | 45226 | 4 | US-09-949-016-12428 |
| 16 | 39 | 10.4 | 30635 | 4 | US-09-949-016-16501 |
| 17 | 39 | 10.4 | 43414 | 4 | US-09-949-016-12839 |
| 18 | 39 | 10.4 | 43415 | 4 | US-09-949-016-16491 |
| 19 | 38.8 | 10.3 | 1941 | 4 | US-09-902-540-4762 |
| 20 | 38.8 | 10.3 | 26492 | 4 | US-09-902-540-1234 |
| 21 | 38.6 | 10.3 | 1940 | 4 | US-09-718-032-1 |
| 22 | 38.6 | 10.3 | 1941 | 3 | US-09-082-737-1 |
| 23 | 38.6 | 10.3 | 2310 | 4 | US-09-774-528-142 |
| 24 | 38.6 | 10.3 | 32495 | 4 | US-09-949-016-14952 |
| 25 | 38.4 | 10.2 | 505 | 4 | US-09-621-376-15639 |
| 26 | 37.8 | 10.1 | 4286 | 4 | US-09-899-634C-1 |
| 27 | 37.8 | 10.1 | 8033 | 4 | US-09-574-779B-134 |

| | | | | | | | |
|---|----|------|------|--------|---|---------------------|-------------------|
| c | 28 | 37.8 | 10.1 | 24942 | 4 | US-09-949-016-13318 | Sequence 13318, A |
| c | 29 | 37.8 | 10.1 | 40237 | 4 | US-09-949-016-16461 | Sequence 16461, A |
| c | 30 | 37.4 | 10.0 | 22712 | 4 | US-09-949-016-16759 | Sequence 16759, A |
| c | 31 | 37.4 | 10.0 | 22712 | 4 | US-09-949-016-16760 | Sequence 16760, A |
| c | 32 | 37.4 | 10.0 | 22712 | 4 | US-09-949-016-16761 | Sequence 16761, A |
| c | 33 | 37.4 | 10.0 | 22712 | 4 | US-09-949-016-16762 | Sequence 16762, A |
| c | 34 | 37.2 | 9.9 | 152331 | 3 | US-09-128-155-16 | Sequence 16, Appl |
| c | 35 | 37 | 9.9 | 2694 | 4 | US-09-902-540-4257 | Sequence 4257, Ap |
| c | 36 | 37 | 9.9 | 24791 | 4 | US-09-902-540-1211 | Sequence 1211, Ap |
| c | 37 | 36.8 | 9.8 | 10029 | 4 | US-09-949-016-11846 | Sequence 11846, A |
| c | 38 | 36.8 | 9.8 | 10029 | 4 | US-09-949-016-16140 | Sequence 16140, A |
| c | 39 | 36.6 | 9.8 | 624 | 4 | US-09-902-540-3091 | Sequence 3091, Ap |
| c | 40 | 36.6 | 9.8 | 15789 | 4 | US-09-902-540-1139 | Sequence 1139, Ap |
| c | 41 | 36.6 | 9.8 | 31199 | 4 | US-09-949-016-16516 | Sequence 16516, A |
| c | 42 | 36.4 | 9.7 | 187595 | 4 | US-09-949-016-15546 | Sequence 15546, A |
| c | 43 | 36.2 | 9.7 | 2608 | 4 | US-09-904-615-16 | Sequence 16, Appl |
| c | 44 | 36.2 | 9.7 | 18112 | 4 | US-09-949-016-13648 | Sequence 13648, A |
| c | 45 | 36.2 | 9.7 | 25419 | 4 | US-09-949-016-15476 | Sequence 15476, A |

ALIGNMENTS

RESULT 1

PCT-US93-06251-23
; Sequence 23, Application PC/TUS9306251
; GENERAL INFORMATION:
; APPLICANT: Wickstrom, Eric and Rife, Jason P.
; TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing Stereospecific Alkylphosphonates and Arylphosphonates
; NUMBER OF SEQUENCES: 93
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: NY
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/06251
; FILING DATE: 19930630
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8586
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-742-4343
; TELEFAX: 516-742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11558 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
PCT-US93-06251-23

Query Match 11.9%; Score 44.8; DB 5; Length 11558;
Best Local Similarity 56.7%; Pred. No. 0.025;
Matches 102; Conservative 0; Mismatches 77; Indels 1; Gaps 1;
QY 74 GGCGTGGCGGGGAGCGGACTGTCGTCGTGCTGAGCGCGGCGGAGCGGCGGAGC 133
Db 428 GGGAGGGGGGGTATGGGGACCTTCAGACGACGAGGGGCTTCGGGGGCGTGGGGCGCC 487
QY 134 GGCTGATCGGCTCCCTCGAACTGGGAGGTCCAGTGGGGTCG-CTTAGGGCCCAAGCCC 192

```
Db 488 TCGGAGCGAGCGCGCGGTGACGCGGTGCGCTGGCGCGAGTGATGACAGCGGCGC 547
QY 193 CCACCGCGGTCCAAAGCTCCAGGCGCTCCCGAGGACACGGGTGCTCGGCGCTTCCTTCG 252
Db 548 CCGGCGCGAAGCCCGAGCGCGCGGGCTCCCGACCGCGGCGCTCCCGCGCTCCCGCG 607

RESULT 2
US-09-949-016-15003/c
; Sequence 15003, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15003
; LENGTH: 9622
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15003

Query Match 11.2%; Score 42; DB 4; Length 9622;
Best Local Similarity 49.5%; Pred. No. 0.13;
Matches 108; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

QY 38 GAGCGAGGCGGCGCGAGTGAGCTGAGTCCGGGTGCGGCGGAGCGGACTGTC 97
Db 7138 GCGTGAGGCGCTGCTGGAAGGGGTGCTGCTGCCGTAGGGGCTCTGTTCCATGGCCACA 7079

QY 98 CGTGTGCTGAGCGCGCGGAGAGCGGCGCGAGGCGGTGATCGGCTCCCTCGAACTGG 157
Db 7078 TCTGCTGCTGTGTGGGGCGAGCGGCGGTAGGAGCCATCATGCTCCATGCGGCTGG 7019

QY 158 GGAGGTCCAGTGGGTGCTTGGGCGCCAAAGCCCGCTCCAAAGCTCCAGG 217
Db 7018 ACAGGACCTCTGGGGCGACAGTGAGCTTACCCCTGGCTGGTCCACGCGCTCCAGC 6959

QY 218 GCTCCCGAGGACCGGTGCTCGGCGCTTCTCTCGTGC 255
Db 6958 TACACCCGAGCGCTCCCTTCTGATCTCTTGGGAC 6921

RESULT 3
US-09-902-540-4944
; Sequence 4944, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; LENGTH: 6651
; TYPE: DNA
US-09-949-016-15003

Query Match 11.2%; Score 42; DB 4; Length 9622;
Best Local Similarity 49.5%; Pred. No. 0.13;
Matches 108; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

QY 38 GAGCGAGGCGGCGCGAGTGAGCTGAGTCCGGGTGCGGCGGAGCGGACTGTC 97
Db 7138 GCGTGAGGCGCTGCTGGAAGGGGTGCTGCTGCCGTAGGGGCTCTGTTCCATGGCCACA 7079

QY 98 CGTGTGCTGAGCGCGCGGAGAGCGGCGCGAGGCGGTGATCGGCTCCCTCGAACTGG 157
Db 7078 TCTGCTGCTGTGTGGGGCGAGCGGCGGTAGGAGCCATCATGCTCCATGCGGCTGG 7019

QY 158 GGAGGTCCAGTGGGTGCTTGGGCGCCAAAGCCCGCTCCAAAGCTCCAGG 217
Db 7018 ACAGGACCTCTGGGGCGACAGTGAGCTTACCCCTGGCTGGTCCACGCGCTCCAGC 6959

QY 218 GCTCCCGAGGACCGGTGCTCGGCGCTTCTCTCGTGC 255
Db 6958 TACACCCGAGCGCTCCCTTCTGATCTCTTGGGAC 6921

RESULT 4
US-09-902-540-1243
; Sequence 1243, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1243
; LENGTH: 30780
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(30780)
; OTHER INFORMATION: unsure at all n locations
US-09-902-540-1243

Query Match 10.8%; Score 40.4; DB 4; Length 30780;
Best Local Similarity 51.3%; Pred. No. 0.43;
Matches 120; Conservative 0; Mismatches 111; Indels 3; Gaps 1;

QY 39 AGCGAGGCGAGGCGCGAGTGGAAGCTGAGTTCGGGGTGGCGGGAGGCGACTGTCC 98
Db 28499 ATCGAGGCGAGGCGCGCTCGAGCCCGGTGTCGGGACACGCGCATGAAGCACGCGGCC 28558

QY 99 GTGGTCTGAGCGCGCGGAGAGCGGCGCGAGCGGCTGATCGGCTCCCTCGAACTGG 158
Db 28559 CTGGAGCTG---CTCTTCTGACCGCCGTCGCGGTATGAGCGCTGTCTGTGGCGAG 28615

QY 159 GAGGTCCAGTGGGTGCTTAGGGCCCAAGCCCGCTCCAAAGCTCCAGG 218
Db 28616 GTGGAGCTGGTGAAGTCTGTGCAATTGGCGCTGTGGGCGCCACGCGGTGCTGATCCAATG 28675

QY 219 CTCTCCCGAGGACCGGTGCTCGGCGCTTCTTCTCGGTGAGAAAGTCCGCCCTGG 272
Db 28676 TCGGAGCTGTTTCGCGGGGATGCTCAAGTCCATCGGCGGGAAGTGCCTGCCCGG 28729

RESULT 5
US-09-949-016-1335
```

```
; ORGANISM: Myxococcus xanthus
US-09-902-540-4944

Query Match 10.8%; Score 40.4; DB 4; Length 6651;
Best Local Similarity 51.3%; Pred. No. 0.32;
Matches 120; Conservative 0; Mismatches 111; Indels 3; Gaps 1;

QY 39 AGCGAGGCGAGGCGCGAGTGGAAGCTGAGTTCGGGGTGGCGGGAGGCGACTGTCC 98
Db 4498 ATCGAGGCGAGGCGCGCTCGAGCCCGGTGTCGGGACACGCGCATGAAGCACGCGGCC 4557

QY 99 GTGGTCTGAGCGCGCGGAGAGCGGCGCGAGCGGCTGATCGGCTCCCTCGAACTGG 158
Db 4558 CTGGAGCTG---CTCTTCTGACCGCCGTCGCGGTATGAGCGCTGTCTGTGGCGAG 4614

QY 159 GAGGTCCAGTGGGTGCTTAGGGCCCAAGCCCGCTCCAAAGCTCCAGG 218
Db 4615 GTGGAGCTGGTGAAGTCTGTGCAATTGCGCTGTGGGCGCCACGCGGTGCTGATCCAATG 4674

QY 219 CTCTCCCGAGGACCGGTGCTCGGCGCTTCTTCTCGGTGAGAAAGTCCGCCCTGG 272
Db 4675 TCGGAGCTGTTTCGCGGGGATGCTCAAGTCCATCGGCGGGAAGTGCCTGCCCGG 4728

RESULT 4
US-09-902-540-1243
; Sequence 1243, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1243
; LENGTH: 30780
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(30780)
; OTHER INFORMATION: unsure at all n locations
US-09-902-540-1243

Query Match 10.8%; Score 40.4; DB 4; Length 30780;
Best Local Similarity 51.3%; Pred. No. 0.43;
Matches 120; Conservative 0; Mismatches 111; Indels 3; Gaps 1;

QY 39 AGCGAGGCGAGGCGCGAGTGGAAGCTGAGTTCGGGGTGGCGGGAGGCGACTGTCC 98
Db 28499 ATCGAGGCGAGGCGCGCTCGAGCCCGGTGTCGGGACACGCGCATGAAGCACGCGGCC 28558

QY 99 GTGGTCTGAGCGCGCGGAGAGCGGCGCGAGCGGCTGATCGGCTCCCTCGAACTGG 158
Db 28559 CTGGAGCTG---CTCTTCTGACCGCCGTCGCGGTATGAGCGCTGTCTGTGGCGAG 28615

QY 159 GAGGTCCAGTGGGTGCTTAGGGCCCAAGCCCGCTCCAAAGCTCCAGG 218
Db 28616 GTGGAGCTGGTGAAGTCTGTGCAATTGGCGCTGTGGGCGCCACGCGGTGCTGATCCAATG 28675

QY 219 CTCTCCCGAGGACCGGTGCTCGGCGCTTCTTCTCGGTGAGAAAGTCCGCCCTGG 272
Db 28676 TCGGAGCTGTTTCGCGGGGATGCTCAAGTCCATCGGCGGGAAGTGCCTGCCCGG 28729

RESULT 5
US-09-949-016-1335
```


; Sequence 1335, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1335
; LENGTH: 2802
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1335

Query Match 10.7%; Score 40.2; DB 4; Length 2802;
Best Local Similarity 49.8%; Pred. No. 0.31;
Matches 102; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 23 ACCCTGGGCGCAACGAGCGAGCGGGCGCGAGTGGAGCTGGAGTTCGGGGTGGGC 82
DB 569 ACCACGCCAGAGGGGCGCCAGGAGGAGCGAGCGCGGGCGCGAGTTCGGGGTGGGC 628

QY 83 GGGGAGGCGCACTCTCCGTGTGTGTCAGCGCGCGAGCGGGCGCGAGTTCGGGGTGGGC 142
DB 629 GAGCGGGTGGCGCAGTGTGTGACAGCGCGAGCGGGCGCGAGTTCGGGGTGGGC 688

QY 143 GCTCCCTCGAACTGGGAGTGCAGTGGGTGCGTTAGGGCCCAAGAGCCCCACCGGCT 202
DB 689 TCCAGGAGGCGCTCAGGGGTCCCGAGGAGTCTCCCGGAGCAAGCGCCCTCTCCGGG 748

QY 203 CCAAAAGCTCCAGGGCCTCCCGAG 227
DB 749 CCTGATGTCGGCACCCCCAGCCTG 773

RESULT 6
US-09-688-188B-102
; Sequence 102, Application US/09688188B
; Patent No. 6656716
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHYTE, DAVID
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES
; CURRENT APPLICATION NUMBER: US/09/688,188B
; CURRENT FILING DATE: 2000-10-16
; PRIOR FILING DATE: 09/291,417
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: 60/081,784
; PRIOR FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 102
; LENGTH: 2806
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-688-188B-102

Query Match 10.7%; Score 40.2; DB 4; Length 2806;
Best Local Similarity 49.8%; Pred. No. 0.31;
Matches 102; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 23 ACCCTGGGCGCAACGAGCGAGCGGGCGCGAGTGGAGTTCGGGGTGGGC 82
DB 567 ACCACGCCAGAGGGGCGCCAGGAGGAGCGAGCGGGCGGTTCGGGGTTCACAGC 626

QY 83 GGGGAGGCGCACTGTCTCGTGTGTGTCAGCGCGCGAGAGCGGGCGCGAGTTCGGGGTGGGC 142
DB 627 GAGCGGGTGGCGGCGAGTGTGTGACAGCGCGAGCGGGCGCGAGTTCGGGGTTCACAGC 686

QY 143 GCTCCCTCGAACTGGGAGTGCAGTGGGTGCGTTAGGGCCCAAGAGCCCCACCGGCT 202
DB 687 TCCAGGAGGCGTCCAGGGGTCCCGAGGAGTCTCCCGGAGCAAGCGCCCTCTCCGGG 746

QY 203 CCAAAAGCTCCAGGGCCTCCCGAG 227
DB 747 CCTGATGTCGGCACCCCCAGCCTG 771

RESULT 7
US-09-291-417D-102
; Sequence 102, Application US/09291417D
; Patent No. 6680170
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHYTE, DAVID
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES
; FILE REFERENCE: 038602/0329
; CURRENT APPLICATION NUMBER: US/09/291,417D
; CURRENT FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 60/081,784
; PRIOR FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 102
; LENGTH: 2806
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-291-417D-102

Query Match 10.7%; Score 40.2; DB 4; Length 2806;
Best Local Similarity 49.8%; Pred. No. 0.31;
Matches 102; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 23 ACCCTGGGCGCAACGAGCGAGCGGGCGCGAGTGGAGTTCGGGGTGGGC 82
DB 567 ACCACGCCAGAGGGGCGCCAGGAGGAGCGAGCGGGCGGTTCGGGGTTCACAGC 626

QY 83 GGGGAGGCGCACTGTCTCGTGTGTGTCAGCGCGCGAGAGCGGGCGCGAGTTCGGGGTGGGC 142
DB 627 GAGCGGGTGGCGGCGAGTGTGTGACAGCGCGAGCGGGCGCGAGTTCGGGGTTCACAGC 686

QY 143 GCTCCCTCGAACTGGGAGTGCAGTGGGTGCGTTAGGGCCCAAGAGCCCCACCGGCT 202
DB 687 TCCAGGAGGCGTCCAGGGGTCCCGAGGAGTCTCCCGGAGCAAGCGCCCTCTCCGGG 746

QY 203 CCAAAAGCTCCAGGGCCTCCCGAG 227
DB 747 CCTGATGTCGGCACCCCCAGCCTG 771

RESULT 8
US-09-949-016-794
; Sequence 794, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768

Db 567 ACCACGCCAGAGGGGCGCCAGGAGGAGCGAGCGGGCGGTTCGGGGTTCACAGC 626

QY 83 GGGGAGGCGCACTGTCTCGTGTGTGTCAGCGCGCGAGAGCGGGCGCGAGTTCGGGGTGGGC 142

Db 627 GAGCGGGTGGCGGCGAGTGTGTGACAGCGCGAGCGGGCGCGAGAGAGGCGCCCAAGTCT 686

QY 143 GCTCCCTCGAACTGGGAGTGCAGTGGGTGCGTTAGGGCCCAAGAGCCCCACCGGCT 202

Db 687 TCCAGGAGGCGTCCAGGGGTCCCGAGGAGTCTCCCGGAGCAAGCGCCCTCTCCGGG 746

QY 203 CCAAAAGCTCCAGGGCCTCCCGAG 227

Db 747 CCTGATGTCGGCACCCCCAGCCTG 771

RESULT 7
US-09-291-417D-102
; Sequence 102, Application US/09291417D
; Patent No. 6680170
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: MARTINEZ, RICARDO
; APPLICANT: WHYTE, DAVID
; TITLE OF INVENTION: STE20-RELATED PROTEIN KINASES
; FILE REFERENCE: 038602/0329
; CURRENT APPLICATION NUMBER: US/09/291,417D
; CURRENT FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 60/081,784
; PRIOR FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 102
; LENGTH: 2806
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-291-417D-102

Query Match 10.7%; Score 40.2; DB 4; Length 2806;
Best Local Similarity 49.8%; Pred. No. 0.31;
Matches 102; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 23 ACCCTGGGCGCAACGAGCGAGCGGGCGCGAGTGGAGTTCGGGGTGGGC 82
DB 567 ACCACGCCAGAGGGGCGCCAGGAGGAGCGAGCGGGCGGTTCGGGGTTCACAGC 626

QY 83 GGGGAGGCGCACTGTCTCGTGTGTGTCAGCGCGCGAGAGCGGGCGCGAGTTCGGGGTGGGC 142
DB 627 GAGCGGGTGGCGGCGAGTGTGTGACAGCGCGAGCGGGCGCGAGTTCGGGGTTCACAGC 686

QY 143 GCTCCCTCGAACTGGGAGTGCAGTGGGTGCGTTAGGGCCCAAGAGCCCCACCGGCT 202
DB 687 TCCAGGAGGCGTCCAGGGGTCCCGAGGAGTCTCCCGGAGCAAGCGCCCTCTCCGGG 746

QY 203 CCAAAAGCTCCAGGGCCTCCCGAG 227

Db 747 CCTGATGTCGGCACCCCCAGCCTG 771

RESULT 8
US-09-949-016-794
; Sequence 794, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768

Query Match 10.7%; Score 40; DB 4; Length 21295;
Best Local Similarity 47.9%; Pred. No. 0.51; Mismatches 125; Indels 0; Gaps 0;
Matches 115; Conservative 0;

QY 45 GCAGGGCGGAGTGAAGCTGGAGTTCCGGGGTGGGGGAGGCGACTGTCCGTGGTG 104
DB 16255 GGAGGAGAAAGCGCGCTGGCAGCGCATCCAGAGCTGCTCGAGGGGCGGGGCGCGG 16314

QY 105 CTGAGCCCGCGAGAGCGGGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGAGGTC 164
DB 16315 CGCAGCGCGCCACGCGCGCGCGGGAAGCGCCCGCGGACACCCCGGCTGGCGCGGG 16374

QY 165 CAGTGGGGTGGCTTAGGGCCCAAGCCCGCCACCGGCTCCAAAAGCTCCAGGGCCTCCC 224
DB 16375 CTCGGCGGACGAGCTGCACGAGCGGCGCCCTTTCCGCAACGGGCTCACGGTGGGACGCT 16434

QY 225 CAGCAGCGGTGCTCGGCCCTTCCTTGGTTCAGAAAGTCCGCCCTGGGGGCGAGTTCGTC 284
DB 16435 CAGGGGCGCGCTGAGCTGATCCACGCGTCGAAGGTGGCGGCTTTGTAGGCTCCGAC 16494

RESULT 12

US-09-949-016-124631
; Sequence 124631, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 124631
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-124631

Query Match 10.5%; Score 39.2; DB 4; Length 601;
Best Local Similarity 48.2%; Pred. No. 0.42; Mismatches 118; Indels 0; Gaps 0;
Matches 110; Conservative 0;

QY 49 GGGCGGAGTGAAGCTGGAGTTCCGGGGTGGGGGAGGCGACTGTCCGTGGTCTGA 108
DB 84 GGAGTGGTGGTCCCAAGGGGGTCTGTGGCGGTGGAGAGCGTGTCCCTCGGACTCG 143

QY 109 GCGCCCGCAGAGCGGGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGGAGGTCCAGT 168
DB 144 GCGCGCCCGACGCTGGCGGCTCCCGCTGAGCGGCTGCCGCTGCCGCGACCACTGCT 203

QY 169 GGGGTGCTTAGGGCCCAAGCCCCACCGGCTCCAAAAGCTCCAGGGCCTCCCGAGG 228
DB 204 CACCTCCCGGGGGCTCGCCAGGCGGCGCTGCGGTGGCGGCTGCCCGGCTGCCCAGG 263

QY 229 CACCGGTGCTCGGCCCTTCCTTCGCTCAGAAAGTCCGCCCTCGGGGC 276
DB 264 GAGCGGAGTGGGTTTCAGCGGGCGGCTGCGCACTCGGCTGGGC 311

RESULT 13

US-09-949-016-15257
; Sequence 15257, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.

Query Match 10.5%; Score 39.2; DB 4; Length 45225;
Best Local Similarity 53.2%; Pred. No. 0.97; Mismatches 83; Indels 0; Gaps 0;
Matches 83; Conservative 0;

QY 28 GGGCGCAACGAGCGAGGCGGCGGAGTGAAGTTCGGGGTGGGGGCGGGA 87
DB 1994 GCGGCGCGGGCGGCGGCAAGCGCGTGGCGGGCTCGGCTGCGGAGTGGGGCGGG 1935

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15257
; LENGTH: 8374
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15257

Query Match 10.5%; Score 39.2; DB 4; Length 8374;
Best Local Similarity 48.2%; Pred. No. 0.7; Mismatches 118; Indels 0; Gaps 0;
Matches 110; Conservative 0;

QY 49 GGGCGGAGTGAAGCTGGAGTTCCGGGGTGGGGGAGGCGACTGTCCGTGGTCTGA 108
DB 5933 GGGAGTGGTGGTCCCAAGGGGGTCTGTGGCGGTGGAGAGCGTGTCCCTCGGACTCG 5992

QY 109 GCGCCCGCAGAGCGGGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGGAGGTCCAGT 168
DB 5993 GCGCGCCCGACGCTGCGGGTCCCGCTGAGCGGCTGCCGCTGCCGCGGACCACTGCT 6052

QY 169 GGGGTGCTTAGGGCCCAAGCCCCACCGGCTCCAAAAGCTCCAGGGCCTCCCGAGG 228
DB 6053 CACCTCCCGGGGGCTCGCCAGGCGGCGCTGCGTGGCGGGTGGCGGCTGCCCAGG 6112

QY 229 CACCGGTGCTCGGCCCTTCCTTCGCTCAGAAAGTCCGCCCTCGGGGC 276
DB 6113 GAGCGGAGTGGGTTTCAGCGGGCGGCTGCGCACTCGGCTGGGC 6160

RESULT 14

US-09-949-016-12428/c
; Sequence 12428, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12428
; LENGTH: 45225
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12428

Query Match 10.5%; Score 39.2; DB 4; Length 45225;
Best Local Similarity 53.2%; Pred. No. 0.97; Mismatches 83; Indels 0; Gaps 0;
Matches 83; Conservative 0;

QY 88 GCGGACTGTCCTGTGCTGAGCGCGGAGAGCGGCGCGGAGCGGCGGCTGATCGGCTCC 147
 Db 1934 CGCGGAGTGAGTGGAGAGAGTGCGCGGTGTCGCGCCCGCGGCTGCGCGGCGGCT 1875
 QY 148 CTCGAACCTGGGAGTCCAGTGGGTGCGCTTAGGGC 183
 Db 1874 CGCCCGGGTGGGGCGGGTGGGGCGGACGGGGC 1839

RESULT 15
 US-09-949-016-13654/c
 ; Sequence 13654, Application US/09949016
 ; Patent No.: 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 13654
 ; LENGTH: 45226
 ; TYPE: DNA
 ; ORGANISM: Human
 ; US-09-949-016-13654

Query Match 10.5%; Score 39.2; DB 4; Length 45226;
 Best Local Similarity 53.2%; Pred. No. 0.97;
 Matches 83; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
 QY 28 GGGCGCAACGAGCGGAGCGGCGGAGTGGAGCTGCGGCTCGGGGCGGAGTGGGCGGGA 87
 Db 1994 GCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1935
 QY 88 GCGGACTGTCCTGTGCTGAGCGCGGAGAGCGGCGGCGGCGGCTGATCGGCTCC 147
 Db 1934 CGCGGAGTGAGTGGAGAGAGTGCGCGGTGTCGCGCCCGCGGCTGCGGGCGGCT 1875
 QY 148 CTCGAACCTGGGAGTCCAGTGGGTGCGCTTAGGGC 183
 Db 1874 CGCCCGGGTGGGGCGGGTGGGGCGGACGGGGC 1839

Search completed: April 7, 2005, 15:00:31
 Job time : 88.5385 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2005, 10:54:05 ; Search time 335.745 Seconds
(without alignments)
6768.525 Million cell updates/sec

Title: US-09-977-579-4_COPY_1_375

Perfect score: 375

Sequence: 1 cctcccttcgagctgagc.....gaagccgcccagccagagaag 375

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5615251 seqs, 3030001701 residues

Total number of hits satisfying chosen parameters: 11230502

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
22: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description | |
|------------|-------|-------------|--------|-------|---------------------|----------------------|
| 1 | 375 | 100.0 | 1261 | 11 | US-09-977-579-4 | |
| c | 2 | 45.2 | 12.1 | 497 | 17 | US-10-424-599-133788 |
| | 3 | 44.8 | 11.9 | 29340 | 18 | US-10-322-281-642 |
| c | 4 | 43 | 396 | 18 | US-10-437-963-98256 | |
| | 5 | 43 | 11.5 | 1021 | 18 | US-10-767-701-11876 |
| c | 6 | 42.6 | 11.4 | 419 | 18 | US-10-425-115-24179 |
| | 7 | 42.2 | 11.3 | 1211 | 18 | US-10-437-963-27585 |
| c | 8 | 41.4 | 11.0 | 696 | 18 | US-10-437-963-90177 |
| | 9 | 41.4 | 11.0 | 58038 | 19 | US-10-741-600-17942 |
| 10 | 41.4 | 11.0 | 333811 | 19 | US-10-741-600-17681 | |
| 11 | 41.2 | 11.0 | 922 | 15 | US-10-017-161-1435 | |

| | | | | | | |
|----|------|------|-------|----|----------------------|--------------------|
| 12 | 41.2 | 11.0 | 922 | 17 | US-10-292-798-1161 | Sequence 1161, Ap |
| 13 | 41 | 10.9 | 1097 | 9 | US-09-771-161A-58 | Sequence 58, Appl |
| 14 | 41 | 10.9 | 3382 | 9 | US-09-771-161A-59 | Sequence 59, Appl |
| 15 | 40.2 | 10.7 | 786 | 17 | US-10-260-238-52 | Sequence 52, Appl |
| 16 | 40.2 | 10.7 | 791 | 18 | US-10-437-963-84059 | Sequence 84059, A |
| 17 | 40.2 | 10.7 | 1071 | 18 | US-10-437-963-84064 | Sequence 84064, A |
| 18 | 40.2 | 10.7 | 2806 | 18 | US-09-291-417-102 | Sequence 102, App |
| 19 | 40.2 | 10.7 | 2806 | 18 | US-10-725-329-102 | Sequence 102, App |
| 20 | 40.2 | 10.7 | 2838 | 16 | US-10-134-102-3 | Sequence 3, Appli |
| 21 | 40.2 | 10.7 | 20000 | 17 | US-10-188-777-11 | Sequence 11, Appli |
| 22 | 40 | 10.7 | 645 | 18 | US-10-767-701-25929 | Sequence 25929, A |
| 23 | 40 | 10.7 | 940 | 18 | US-10-437-963-40850 | Sequence 40850, A |
| 24 | 39.8 | 10.6 | 332 | 9 | US-09-960-352-2238 | Sequence 2238, Ap |
| 25 | 39.8 | 10.6 | 863 | 18 | US-10-437-963-71791 | Sequence 71791, A |
| 26 | 39.6 | 10.6 | 712 | 18 | US-10-425-115-118813 | Sequence 118813, |
| 27 | 39.6 | 10.6 | 735 | 18 | US-10-437-963-84061 | Sequence 84061, A |
| 28 | 39.6 | 10.6 | 1092 | 18 | US-10-767-701-9739 | Sequence 9739, Ap |
| 29 | 39.6 | 10.6 | 4233 | 18 | US-10-723-860-4143 | Sequence 4143, Ap |
| 30 | 39.6 | 10.6 | 4244 | 18 | US-10-723-860-8016 | Sequence 8016, Ap |
| 31 | 39.2 | 10.5 | 759 | 17 | US-10-260-238-230 | Sequence 230, App |
| 32 | 39.2 | 10.5 | 972 | 18 | US-10-425-115-17653 | Sequence 17653, A |
| 33 | 39.2 | 10.5 | 1294 | 18 | US-10-437-963-73166 | Sequence 73166, A |
| 34 | 39.2 | 10.5 | 6888 | 18 | US-10-602-494-20 | Sequence 20, Appl |
| 35 | 39 | 10.4 | 485 | 17 | US-10-424-598-65670 | Sequence 65670, A |
| 36 | 39 | 10.4 | 587 | 16 | US-10-029-386-22980 | Sequence 22980, A |
| 37 | 39 | 10.4 | 936 | 9 | US-09-974-300-836 | Sequence 836, App |
| 38 | 39 | 10.4 | 1087 | 17 | US-10-424-599-81404 | Sequence 81404, A |
| 39 | 39 | 10.4 | 88400 | 18 | US-10-844-716-1 | Sequence 1, Appli |
| 40 | 38.8 | 10.3 | 985 | 18 | US-10-739-930-4166 | Sequence 4166, Ap |
| 41 | 38.6 | 10.3 | 263 | 18 | US-10-021-323-16665 | Sequence 16665, A |
| 42 | 38.6 | 10.3 | 550 | 16 | US-10-029-386-24925 | Sequence 24925, A |
| 43 | 38.6 | 10.3 | 1647 | 18 | US-10-437-963-37566 | Sequence 37566, A |
| 44 | 38.6 | 10.3 | 1940 | 17 | US-10-693-367-1 | Sequence 1, Appli |
| 45 | 38.6 | 10.3 | 2310 | 17 | US-10-120-988-142 | Sequence 142, App |

ALIGNMENTS

RESULT 1

US-09-977-579-4
; Sequence 4, Application US/09977579
; Publication No. US20040248240A1
; GENERAL INFORMATION:
; APPLICANT: Cambridge University Technical Services
; TITLE OF INVENTION: A novel family of beta sub-unit proteins from a voltage gated sod
; TITLE OF INVENTION: channel
; FILE REFERENCE: nucleic acids encoding them and therapeutic or diagnostic uses ti
; FILE REFERENCE: 674558-2001
; CURRENT FILING DATE: 2001-10-15
; CURRENT APPLICATION NUMBER: US/09/977,579
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: PCT/EP00/01783
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60,129,473
; PRIOR FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 4
; LENGTH: 1261
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-977-579-4

| | | | | | | | |
|-----------------------|---------|---|--------|------------|-----|--------|-------|
| Query Match | 100.0%; | Score | 375; | DB | 11; | Length | 1261; |
| Best Local Similarity | 100.0%; | Pred. No. | 7e-98; | | | | |
| Matches | 375; | Conservative | 0; | Mismatches | 0; | Indels | 0; |
| Gaps | 0; | | | | | | |
| Qy | 1 | CCCTCCCTTCGAGCTGAGCTTACCTCGGCGCAACGACGAGGCGGCGGCGAGTGG | 60 | | | | |
| Db | 1 | CCCTCCCTTCGAGCTGAGCTTACCTCGGCGCAACGACGAGGCGGCGGCGAGTGG | 60 | | | | |
| Qy | 61 | AAGCTGGAGTCCCGGGTGGCGGGGAGGCGACTGTCCGTGGTCTGAGCCGCCGCGAGA | 120 | | | | |

Db 61 AAGCTGGAGTTCGGGGTGGGGGAGGCGACTGTCCGTGGTGTCTGAGCGCGCGGAGA 120
QY 121 GCGGGCGCGAGCGGTGATCGCTCCCTCGAACTGGGAGGTCCAGTGGGGTTCGCTTAG 180
Db 121 GCGGGCGCGAGCGGTGATCGCTCCCTCGAACTGGGAGGTCCAGTGGGGTTCGCTTAG 180
QY 181 GCGCCAAAGCCCCACCGGCTCCAAAGCTCCCAAGGCTCCCAAGGCTCCCAAGGCTCCG 240
Db 181 GCGCCAAAGCCCCACCGGCTCCAAAGCTCCCAAGGCTCCCAAGGCTCCCAAGGCTCCG 240
QY 241 GCGCTTCCTTCGCTCAGAAAGTCCGCCCCCTGGGGGCGAGTTCGTCCCAAGGGTTTCCTCG 300
Db 241 GCGCTTCCTTCGCTCAGAAAGTCCGCCCCCTGGGGGCGAGTTCGTCCCAAGGGTTTCCTCG 300
QY 301 AAGAATCTGAGAGGGCGAGTCCCTTGACCGAGGGAATCTCTGTGTAGCTTGAAGC 360
Db 301 AAGAATCTGAGAGGGCGAGTCCCTTGACCGAGGGAATCTCTGTGTAGCTTGAAGC 360
QY 361 CGCCAGCCCCAGAG 375
Db 361 CGCCAGCCCCAGAG 375

RESULT 2

US-10-424-599-133788/c
; Sequence 133788, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 133788
; LENGTH: 497
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(497)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_91816C.1
US-10-424-599-133788

Query Match 12.1%; Score 45.2; DB 17; Length 497;
Best Local Similarity 54.6%; Pred. No. 0.0024;
Matches 89; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
QY 38 GAGCGAGGCGAGGCGCGAGTGGAACTCGAGTTCGGGGTGGGGGCGGAGCGGACTGTC 97
Db 301 GGGGGGAGGGGGGGGGGGGGGGGAGGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 242
QY 98 CTTGTGTCTGAGCGCGCGGAGAGCGGGCGCGAGCGGCTGTATCGGCTCCCTCGAACTGG 157
Db 241 CGGGGGGAGGG 182
QY 158 GGAGGTTCAGTGGGGTTCGCTTAGGGGCCCAAGGCCCCACCCGG 200
Db 181 GGGGAGGG 139

RESULT 3

US-10-322-281-642
; Sequence 642, Application US/1032281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris

; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001000
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 642
; LENGTH: 29340
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-322-281-642
Query Match 11.9%; Score 44.8; DB 18; Length 29340;
Best Local Similarity 56.7%; Pred. No. 0.0026;
Matches 102; Conservative 0; Mismatches 77; Indels 1; Gaps 1;
QY 74 GGGGTGGGGGGGAGCGGAGTGTCTGCTGAGCGCGCGGAGCGGCGGCGGAGC 133
Db 9834 GGGAGGGCGGTAGGGGACCTTTCAGAGCCAGGAGGGCTTTCGGGGGGCTGGGGCGGC 9893
QY 134 GGCTGATCGGCTCCCTCGAACTGGGGAGGTCCAGTGGGTGCG-CTTAGGGCCCAAGCCC 192
Db 9894 TCGGAGCGGAGCGCGGCTCGACGCGGTGCGTGGCGGAGTGTATGCAGACGGCGC 9953
QY 193 CCACCCGGCTCCAAAAGCTCCAGGGCTCCCGAGGACACCGGTGCTCGGCCCTTCTTCG 252
Db 9954 CCGGCCCGAACCCCGAGCCCGCGGGGTCTCCCGACCGCGGCTCCCGCCCTCCCGCG 10013
RESULT 4
US-10-437-963-98256
; Sequence 98256, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrei A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 98256
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_9617C.1
US-10-437-963-98256

Query Match 11.5%; Score 43; DB 18; Length 396;
Best Local Similarity 50.7%; Pred. No. 0.011;
Matches 103; Conservative 0; Mismatches 100; Indels 0; Gaps 0;
QY 72 CCGGGGTGGGGGAGGCGGACTGTCCGTGGTGTGAGCGCGCGGAGAGCGGGCGCGA 131
Db 129 CGCGCGGCGAGGAGGCGGCGGCTGTGGCCGGGGGGCGCGGAGCGCGGGCAGG 188
QY 132 GCGGCTGATCGGCTCCCTCGAACTGGGAGGTCCAGTGGGGTTCGCTTAGGGGCCCAAGCC 191
Db 189 CGGGGTGGCGGTGCGCGGGCGGTCCCGAGCAGGCTTGGCGGCGCTGTGCGCGGGAC 248
QY 192 CCCACCGGCTCCAAAAGCTCCAGGGCTCCCGAGGACCGGTGCTCGGCCCTTCTTCG 251
Db 249 CGCGCGCTGCGCGCGGCTGCGCGGGCGCGGCCCGCCGACGAGCGGCGAGCGCGCGG 308

Qy 252 GGTCAGAAAGTCGCCCTGGGG 274
Db 309 GCTGGCAGGGCCGCCCGCGG 331

RESULT 5

```

US-10-767-701-11876/c
; Sequence 11876, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecule
; FILE OF INVENTION: Plants and Uses Thereof
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128

```

```

; SEQ ID NO 11876
;
; LENGTH: 1021
;
; TYPE: DNA
;
; ORGANISM: Sorghum bicolor
;
; FEATURE:
;
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS31593_1
US-10-767-703-11876

```

Query Match 11.5%; Score 43; DB 18; Length 1021;
Best Local Similarity 54.8%; Pred. NO. 0.01;
Matches 85; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

| | | | |
|----|-----|--|-----|
| Qy | 28 | GGGCGCAACGAGCGAGGCGAGGGCGCGAGCTGGAGTTCCGGGCTGGCGCGGGA | 87 |
| | | | |
| Db | 232 | GGGCGGCGAGGACCGAGGAGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG | 173 |
| | | | |
| Qy | 88 | GGGCACTGTCGTGTGCTGACCGCGGGCGAGAGCGGGCGCGGAGCGGCTGATCGGCTCC | 147 |
| | | | |
| Db | 172 | GG | 113 |
| | | | |
| Qy | 148 | CTCGAACTGGGAGGTCCAGTCGGGTCGCTTAGGG | 182 |
| | | | |
| Db | 112 | GGGGGGGGGGGGAGGCCCGAGCGGCGGGGAGGGGGGGGGGGGGGGGGGGGGGG | 78 |
| | | | |

RESULT 6

```

RESULTS 6
US-10-425-115-24179/c
; Sequence 24179, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326

```

SEQ ID NO 24179

```
; LENGTH: 419
; TYPE: DNA
```

ORGANISM: *Zea mays*

FEATURE:

NAME/KEY: unsure

LOCATION: (1) : (419)

OTHER INFORMATION: u

FEATURE:

OTHER INFORMATION: C

S-10-425-115-24179

Query Match

Best Local Similarity

| | Matches | 119; | Conservative | 0; | Mismatches | 110; | Indels | 1; | Gaps | 1 |
|----|---------|----------------------|---------------------|----------------|-------------------|------|--------|----|------|---|
| Qy | 56 | AGTGGAAAGCTGGAGTTC | CGGGGGTGGGGAGGCGGAC | TGTCCGTGGTGC | TGAGCGCCGG | 115 | | | | |
| Db | 371 | AGTCTCAGCGCGGACGAGT | GGGGTGGCGTCTGCGT | CTCTTGC | GACGGGGGGGGGGG | 312 | | | | |
| Qy | 116 | CGAGAGCGGGCGGAGCGGCT | GTATCGGCTCCCTCGAACT | CGGGAGGTC | TCAGTGGGTGCG | 175 | | | | |
| Db | 311 | GGAGACCGAGCGGAAGCGG | CGGGGACGACCACTGT | CGTGGG | CAGGGCCACGCCCGCGC | 252 | | | | |
| Qy | 176 | CTTAGGGCCCAAGCCCCC | ACCCGCTCCAAAGCTCC | CAAGGGGCT | TCCCAAGGACACCGGT | 235 | | | | |
| Db | 251 | TGGAGGCCCGCCAGCGC | AGGCTCGAGAGCACCGG | AGCCGGCCCCCGG | AGGGGCGACGA | 192 | | | | |
| Qy | 236 | GCTCGGGCCCT-TCTTC | CGGTCAAGAAAGT | CGCCGCCCTCGGGG | CGACGTCTGTC | 284 | | | | |
| Db | 191 | GCTTAGCGAAGTCCGG | CGGCGCAGTAGAC | TTCGGCTT | GAGTGGCGGCC | 142 | | | | |

RESULT 7

```

US-10-437-963-27585/c
; Sequence 27585, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 27585
; LENGTH: 1211
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_32265C.1
US-10-437-963-27585

```

Query Match

| | | | | | |
|----|-----------------------|---|------------------|--------|----------------------------------|
| | Query Match | 51.3%; | Score 42.4; | DB 18; | Length 1411; |
| | Best Local Similarity | 55.9%; | Pred. No. 0.017; | | |
| | Matches | 80; | Conservative | 0; | Mismatches 63; Indels 0; Gaps 0; |
| Qy | 40 | GCAGGACGGGGCGCAGCTGGAACTTCGGGTGGGGAGGGACTGTCCG | 99 | | |
| Dd | 204 | GCCTGGGGGGCGGGGGGGCGCGCGGGGGCGGGGGGGGGGGCGCGC | 145 | | |
| Qy | 100 | TGTTGCTGAGCGCGGGGAGAGCGGGCGCGAGCGGCTGATCGGCTCCCTCGAATGGGG | 159 | | |
| Dd | 144 | GCGCGCGCGCCCCCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG | 85 | | |
| Qy | 160 | AGTTCAGTGGGTTCGCTTAGGG | 182 | | |
| Dd | 84 | GGCGGGCGGGGGCGCGGGGGG | 62 | | |

RESULT 8

RES-101-8
US-143-963-90177/C
; Sequence 90177, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei

APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 90177
LENGTH: 696
TYPE: DNA
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_88871C.1
US-10-437-963-90177

Query Match 11.0%; Score 41.4; DB 18; Length 696;
Best Local Similarity 48.5%; Pred. No. 0.029;
Matches 114; Conservative 0; Mismatches 121; Indels 0; Gaps 0;
QY 10 CCGAGCTGAGCTTACCTCTGGCGCAAAACGAGCGAGCGCGCGCGAGTGGAGCTGGAG 69
DB 594 CGCGCGCGCGACACGCGCGGGATCGGAGCGCGAGCGCGCGCGCGCGCGCGCG 535
QY 70 TTCGGGGTGGCGGGGAGCGACTGTCTCGTGTGTGAGCGCGCGCGAGAGCGCGCG 129
DB 534 GTGAGCGCAGCGCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 475
QY 130 GAGCGCTGATCGGCTCCCTCGAACTGGGGAGTCCAGTGGGTGCTTAGGGCCCAAAG 189
DB 474 GAGGCGCAACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 415
QY 190 CCCCACCGCGCTCCAAAGTCTCCAGGCGCTCCCGAGCACCGGTGTCTCGGCC 244
DB 414 TCCCTCTCTCAGCGCGGAGCGCGCGCGCTCAACCGCGCGCGCGCGCGCGCTCC 360

RESULT 9
US-10-741-600-17942
Sequence 17942, Application US/10741600
Publication No. US20050026169A1
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001499
CURRENT APPLICATION NUMBER: US/10/741,600
CURRENT FILING DATE: 2003-12-22
NUMBER OF SEQ ID NOS: 73997
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 17942
LENGTH: 58038
TYPE: DNA
ORGANISM: Homo sapiens
US-10-741-600-17942

Query Match 11.0%; Score 41.4; DB 19; Length 58038;
Best Local Similarity 49.8%; Pred. No. 0.023;
Matches 132; Conservative 1; Mismatches 127; Indels 5; Gaps 1;
QY 28 GGGCGCAACGAGCGAGCGCGCGCGCGAGTGGAGCTGCGGCTGATCGGCTCC 87
DB 6432 GGACRACACGGGCGCTGGCGACCGCGCGCGAGTGTCCCGAGGCGCGGCTTCGGGGCGGG 6491
QY 88 GCGACTGTCCGTGTGCTGAGCGCGCGCGAGCGCGCGCGAGCGGCTGATCGGCTCC 147
DB 6492 GCCGGTGTCCCTTGGCGGGGCTCAGGGGTGGAGCGCGTGGAGCGCTGTGTTTCG- 6550
QY 148 CTCGAACCTGGGAGGTCCAGTGGGTGCTTTAGGGCCCAAAGCCCGCCCGCTCCAAA 207
DB 6551 ----ACCCCTCGCGCCCTTGGCGCGCGCGCGCGCGCGCTTTACGCGACCTCCCCACA 6606

QY 208 AGCTCCCGGGCTCCCGAGCGCGGTGCTCGGCCCTTCTTCGTCAGAAAGTCGCC 267
DB 6607 AGACTCAAAGGTGCGGGCGGCACTGCTCTCCGCCCTTCCCGGAATCAAACCTTCCCTC 6666
QY 268 CCTGGGGCGAGTTCGTCCCAAGGG 292
DB 6667 TGCAGCGCGCGCTTACGGAGGG 6691

RESULT 10
US-10-741-600-17681
Sequence 17681, Application US/10741600
Publication No. US20050026169A1
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001499
CURRENT APPLICATION NUMBER: US/10/741,600
CURRENT FILING DATE: 2003-12-22
NUMBER OF SEQ ID NOS: 73997
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 17681
LENGTH: 333811
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1)...(333811)
OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-;
US-10-741-600-17681

Query Match 11.0%; Score 41.4; DB 19; Length 333811;
Best Local Similarity 49.8%; Pred. No. 0.021;
Matches 132; Conservative 1; Mismatches 127; Indels 5; Gaps 1;
QY 28 GGGCGCAACGAGCGAGCGCGCGCGAGTGGAGCTGCGGCTTCGGGGCGGG 87
DB 191325 GGACRACACGGGCGCTGGGCGACGCGCGGCGAGTGTCCCGAGGCGCGGCTTCGGGGCGGG 191384
QY 88 GCGCACTGTCTCGTGTGCTGAGCGCGCGCGAGCGCGCGCGAGCGGCTGATCGGCTCC 147
DB 191385 GCCGGTGTCCCTTCGCGGGGCTCAGGGGTGGAGCGGTGCGACGCTGTGTTTCG- 191443
QY 148 CTCGAACCTGGGAGGTCCAGTGGGTGCTTTAGGGCCCAAAGCCCGCCCGCTCCAAA 207
DB 191444 ----ACCCCTCGCGCCCTTGGCGCGCGCGCGCGCGGCTTTACGCGACCTCCCCACA 191499
QY 208 AGCTCCCGGGCTCCCGAGCGCGGTGCTCGGCCCTTCTTCGTCAGAAAGTCGCC 267
DB 191500 AGACTCAAAGGTGCGGGCGGCACTGCTCTCCGCCCTTCCCGGAATCAAACCTTCCCTC 191559
QY 268 CCTGGGGCGAGTTCGTCCCAAGGG 292
DB 191560 TGCAGCGCGCGCTTACGGAGGG 191584

RESULT 11
US-10-017-161-1435
Sequence 1435, Application US/10017161
Publication No. US20030143668A1
GENERAL INFORMATION:
APPLICANT: SUWA, MAKIHO
APPLICANT: ASAI, KIYOSHI
APPLICANT: AKIYAMA, YUTAKA
APPLICANT: ABURATANI, HIROYUKI
TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
FILE REFERENCE: 084335/0152
CURRENT APPLICATION NUMBER: US/10/017,161
CURRENT FILING DATE: 2002-12-18
PRIOR APPLICATION NUMBER: JP 2001/246789
PRIOR FILING DATE: 2001-06-18
NUMBER OF SEQ ID NOS: 2430


```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1435
; LENGTH: 922
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(922)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)..(722)
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1)..(26)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (62)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (124)..(125)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (127)..(128)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (131)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (133)..(134)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (138)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (142)..(145)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (147)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (151)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (311)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (322)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (326)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (345)..(349)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (351)..(361)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (415)..(417)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (422)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (438)..(439)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (441)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (443)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (465)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (469)..(471)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (487)..(488)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (490)..(491)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (364)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (366)..(368)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (370)..(380)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (384)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (386)..(388)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (392)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (394)..(396)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (403)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (413)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (415)..(417)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (422)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (438)..(439)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (441)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (443)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (465)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (469)..(471)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (487)..(488)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (490)..(491)
; OTHER INFORMATION: a, t, c, g, unknown or other
```

```

; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (495)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (497)..(498)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (501)..(505)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (519)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (529)..(530)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (540)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (548)..(549)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (564)..(568)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (571)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (573)

Query Match
Best Local Similarity 11.0%; Score 41.2; DB 15; Length 922;
Matches 70; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

QY 12 GAGCTGAGCTTACCTGGCGCAACGACGAGCGAGCGCGGCGGAGTGAAGCTGAGTT 71
DB 363 GNGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 422

QY 72 CCGGGGTGGCGGAGGCGACTGTCTCGTGTGCTCAGCGCGCGGAGCGGCGCGGA 131
DB 423 GGGGGGGGGGGGGGGGGGNNNGNNGNNGGNGGGGGGGGGGGGNNNNNNNNNN 482

QY 132 GCGGCTGATCGGCTCCCTCGAAGTGGGAGGTCCAGTGGGGGTGCTTAGGG 182
DB 483 GGGGNNNGGNNNGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 533

RESULT 12
US-10-292-798-1161
; Sequence 1161, Application US/10292798
; Publication No. US20030235833A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 084335/166
; CURRENT APPLICATION NUMBER: US/10/292,798
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-12-18

; OTHER INFORMATION: a, t, c, g, unknown or other
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1161
; LENGTH: 922
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; LOCATION: source
; FEATURE:
; LOCATION: (1)..(922)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)..(722)
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1)..(26)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (62)..(62)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (124)..(125)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (127)..(128)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (131)..(131)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (133)..(134)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (138)..(138)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (142)..(145)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (147)..(147)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (151)..(151)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (311)..(311)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (322)..(322)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (326)..(326)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (345)..(349)
; OTHER INFORMATION: a, t, c, g, unknown or other
; FEATURE:
; NAME/KEY: modified_base
```

LOCATION: (351)..(361)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (364)..(364)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (366)..(368)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (370)..(380)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (384)..(384)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (386)..(388)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (392)..(392)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (394)..(396)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (403)..(403)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (413)..(413)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (415)..(417)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (422)..(422)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (438)..(439)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (441)..(441)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (443)..(443)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (445)..(445)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (465)..(465)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (469)..(471)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (487)..(488)

OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (490)..(491)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (495)..(495)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (497)..(498)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (501)..(505)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (519)..(519)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (529)..(530)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (540)..(540)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (548)..(549)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (564)..(568)
OTHER INFORMATION: a, t, c, g, unknown or other
FEATURE:
NAME/KEY: modified base
LOCATION: (571)..(571)
OTHER INFORMATION: a, t, c, g, unknown or other

Query Match 11.0%; Score 41.2; DB 17; Length 922;
Best Local Similarity 40.9%; Pred. No. 0.033;
Matches 70; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

Qy 12 GAGCTGAGCTTACCTGGGGCGCAACGAGCGAGCGGGCGCGAGTGGAGTGGAGTT 71
Db 363 GNGNN 422
Qy 72 CCGGGGTGGGGGAGGAGCTGCTCGTGTGCTGAGCGCCGCGAGAGAGAGAGAGAGAG 131
Db 423 GGG 482
Qy 132 GCGGCTGATCGGCTCCCTCGAACTGGGGAGGTCAGTGGGGTGGCTTAGGG 182
Db 483 GGGGNNNGGNGGNN 533

RESULT 13
US-09-771-161A-58
Sequence 58, Application US/09771161A
Patent No. US2002011081A1
GENERAL INFORMATION:
APPLICANT: LEVINE, et al.
TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
FILE REFERENCE: 802620-2005.1
CURRENT APPLICATION NUMBER: US/09/771,161A
CURRENT FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: 09/724,676
PRIOR FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: 136776
PRIOR FILING DATE: 2000-06-15

```
; PRIOR APPLICATION NUMBER: 135619
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 273
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 58
; LENGTH: 1097
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-771-161A-58

Query Match          10.9%; Score 41; DB 9; Length 1097;
Best Local Similarity 52.0%; Pred. No. 0.037;
Matches 92; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 1 CCTCTCCCTTCCGAGCTGAGCTTACCTGGGCGCAACAGCAGAGGAGGGGCGCGAGTGG 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 23 CCCCCGGCTCGGGGCTGTGAGCGGCTCGGGGCGGGGGTGGCGGGTGTGCGCGGGCGG 82
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 61 AAGCTGGAGTTCGGGGTGGGGGAGGCGACTGTCCGTGGTGTGCTGAGCGGCGGCGAGA 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 83 CGACGCTCTCTTCGGCGGCGCGCGCGCGCCATGCTGGGGCGCGCGCTGGGGCG 142
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 121 GCGGGCGCGAGCGGCTGTATCGGCTCCCTCGAACTGGGAGGTCCAGTGGGGTGGCT 177
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 143 GCGGGCGCGAGTGTGCTCCCGGGGCCGCGCTGAGGGGCCCGCGCGCGCGCT 199
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 14
US-09-771-161A-59
; Sequence 59, Application US/09771161A
; Patent No. US20020110811A1
; GENERAL INFORMATION:
; APPLICANT: LEVINE, et al.
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
; FILE REFERENCE: 802620-2005.1
; CURRENT APPLICATION NUMBER: US/09/771,161A
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/724,676
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 136776
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 135619
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 273
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 59
; LENGTH: 3382
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: -
; LOCATION: (1)..(3382)
; OTHER INFORMATION: "n" can be any nucleotide 'a', 'c', 'g' or 't'
US-09-771-161A-59

Query Match          10.9%; Score 41; DB 9; Length 3382;
Best Local Similarity 52.0%; Pred. No. 0.035;
Matches 92; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 1 CCTCTCCCTTCCGAGCTGAGCTTACCTGGGCGCAACAGCAGAGGAGGGGCGCGAGTGG 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 23 CCCCCGGCTCGGGGCTGTGAGCGGCTCGGGGCGGGGGTGGCGGGGCTGTGCGGGCGG 82
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 61 AAGCTGGAGTTCGGGGTGGGGGAGGCGACTGTCCGTGGTGTGCTGAGCGGCGGCGAGA 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 83 CGACGCTCTCTTCGGCGGCGCGCGCGCGCCATGCTGGGGCGCGCGCTGGGGCG 142
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 121 GCGGGCGCGAGCGGCTGTATCGGCTCCCTCGAACTGGGAGGTCCAGTGGGGTGGCT 177
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 143 GCGGGCGCGAGTGTGCTCCCGGGGCCGCGCTGAGGGGCCCGCGCGCGCGCT 199
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 15

```
US-10-260-238-52/c
; Sequence 52, Application US/10260238
; Publication No. US20040016025A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, Paul R.
; APPLICANT: Moughamer, Todd G.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goff, Stephen A.
; APPLICANT: Katagiri, Fumiyaki
; APPLICANT: Krep, Joel
; APPLICANT: Provart, Nicholas
; APPLICANT: Ricke, Darrell
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION
; FILE REFERENCE: 60111-NP
; CURRENT APPLICATION NUMBER: US/10/260,238
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/325,448
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 6077
; SEQ ID NO 52
; LENGTH: 766
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: N region
; LOCATION: (357)..(357)
; OTHER INFORMATION: n = any nucleotide
US-10-260-238-52

Query Match          10.7%; Score 40.2; DB 17; Length 766;
Best Local Similarity 57.6%; Pred. No. 0.065;
Matches 72; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 12 GAGCTGAGCTTACCTGGGCGCAACAGCAGAGGCGAGGCGCGGAGTGGAACTGGAGTT 71
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 686 GAGCGGCGGCGCGGTCGCGGAGCGAGGTGAGCGAGGCGCGACGAGTGGCGCGGT 627
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 72 CCGGGGTGGGCGGGGAGGCGACTGTCCGTGTGTGCTGAGCGCGCGGAGAGCGGGCGCGGA 131
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 626 CCGTCGCGGACGAGGCGGCGACTGTGTCGTGGAGGCGGAGGAGGCGGAGCGCGCGGA 567
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 132 GCGGC 136
   |||
Db 566 TGAGC 562

Search completed: April 7, 2005, 15:19:32
Job time : 337.745 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2005, 08:50:34 ; Search time 2237.21 Seconds
(without alignments)
6380.306 Million cell updates/sec

Title: US-09-977-579-4_COPY_1_375

Perfect score: 375

Sequence: 1 cccctcctccgagctgagc.....gaagccgagccagccagag 375

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hc:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gss1:*

9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|-----------|---------------------|
| 1 | 369.8 | 98.6 | 4052 | 3 | HSM801563 | AL136589 Homo sapi |
| 2 | 346 | 92.3 | 1078 | 1 | AL534136 | AL534136 AL534136 |
| 3 | 279.4 | 74.5 | 582 | 5 | BP202832 | BP202832 BP202832 |
| 4 | 213.8 | 57.0 | 972 | 5 | BQ066875 | BQ066875 AGENCOURT |
| C 5 | 59 | 15.7 | 935 | 9 | CNS006XX | AL066051 Drosophila |
| C 6 | 53.8 | 14.3 | 1101 | 9 | AL063912 | AL063912 Drosophila |
| 7 | 53.2 | 14.2 | 2243 | 9 | AG381986 | AG381986 Mus muscu |
| C 8 | 53 | 14.1 | 693 | 9 | AG060136 | AG060136 Pan trogl |
| C 9 | 51.8 | 13.8 | 935 | 9 | CNS006XX | AL066051 Drosophila |
| C 10 | 51.6 | 13.8 | 925 | 9 | CNS0091P | AL053013 Drosophila |
| 11 | 50.6 | 13.5 | 925 | 9 | CNS0091P | AL053013 Drosophila |
| 12 | 50.4 | 13.4 | 890 | 9 | CL493826 | CL493826 SAIL_586 |
| 13 | 50 | 13.3 | 964 | 9 | AG127812 | AG127812 Pan trogl |
| 14 | 49.8 | 13.3 | 888 | 4 | BG809572 | BG809572 mgct001xa |
| C 15 | 49.6 | 13.2 | 932 | 9 | CNS0072Q | AL066742 Drosophila |
| 16 | 49.4 | 13.2 | 1289 | 4 | BM807701 | BM807701 AGENCOURT |
| 17 | 49.2 | 13.1 | 892 | 9 | CL482374 | CL482374 SAIL_35 F |
| C 18 | 49.2 | 13.1 | 922 | 9 | CL513414 | CL513414 SAIL_875 |
| 19 | 49.2 | 13.1 | 957 | 9 | CL476240 | CL476240 SAIL_251 |
| 20 | 48.8 | 13.0 | 863 | 9 | AG132157 | AG132157 Pan trogl |
| 21 | 48.6 | 13.0 | 970 | 5 | BUS38852 | BUS38852 AGENCOURT |
| C 22 | 48.6 | 13.0 | 1009 | 9 | CNS010EW | AL098882 Drosophila |
| C 23 | 48.4 | 12.9 | 977 | 9 | CL488066 | CL488066 SAIL_503 |
| C 24 | 48.2 | 12.9 | 950 | 9 | CL501906 | CL501906 SAIL_705 |

ALIGNMENTS

RESULT 1
HSM801563
LOCUS HSM801563 4052 bp mRNA linear HTC 22-SEP-2004
DEFINITION Homo sapiens mRNA; cDNA DKFZp761F182 (from clone DKFZp761F182).
ACCESSION AL136589
VERSION AL136589.1 GI:13276680
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4052)
AUTHORS Ortenwaelder B., Obermaier B., Deutschenbaue S., Schaipp A., Mewes H.W., Weil B., Amid C., Oeinger A., Fobo G., Han M. and Wiemann S.
CONSRMT The German CDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@kfz-heidelberg.de; sequenced by Medigenomix (Martinsried/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFZp761F182) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering: http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp761F182 Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.
FEATURES
Location/Qualifiers
1..4052
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFZp761F182"
/db_xref="taxon:9606"
/clone="DKFZp761F182"
/tissue_type="amygdala"
/clone_lib="761 (synonym: hamy2). Vector pSport1; host DH10B; sites NotI + SalI"
/dev_stage="adult"
/note="voltage-gated sodium channel beta-3 subunit"
1..4052
/gene="DKFZp761F182"
804..1451
/gene="DKFZp761F182"
/codon_start=1
/product="hypothetical protein"

```
/protein_id="CAB66524.1"
/db_xref="GI:13276681"
/db_xref="GOA:Q9NY72"
/db_xref="UniProt/Swiss-Prot:Q9NY72"
translation="MPANRLPPLASLVLIVVSVCFVCEVPSETEAVOQNPMLKLR
CISMKREVEATTVFVYRPEGKDLIVYVRNGHVEGSPFGRLQWNGSKDLQD
VSIIVNTLNDLSGLYTCNSREPEFAHRPFVKTRLIPLRVTEAGEDFTSVVSEI
MYYILLVFLTLMLLIEMTYIKRVYSKAEAAQENASDYLAIPSENKENSAPVPEB"
```

source

1. 1078
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODP005YI02"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

Query Match 92.3%; Score 346; DB 1; Length 1078;
Best Local Similarity 98.9%; Pred. No. 3.7e-74;
Matches 369; Conservative 0; Mismatches 1; Indels 3; Gaps 2;

QY 3 CTCCTTCCGAGCTGAGCTTACCTGGCGCGCAACGAGCGAGGCGGGCGCGAGTGGAA 62
DB 431 CTCCTTCCGAGCTGAGCTTACCTGGCGCGCAACGAGCGAGGCGGGCGCGAGTGGAA 490
QY 63 GCTGGAGTTCCGGGTGGCGGGAGGCGACTGTCGGTGGTCTGAGCGCCCGCGAGGC 122
DB 491 GCTGGAGTTCCGGGTGGCGGGAGGCGACTGTCGGTGGTCTGAGCGCCCGCGAGGC 550
QY 123 GGGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGAGGTCAGTGGGGTGCCTTAGGG 182
DB 551 GGGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGAGGTCAGTGGGGTGCCTTAGGG 610
QY 183 CCCAAGCCCCCAGCGGCTCCAAAGCTCCAGGCGCTCCCGAGCGACCGGTGCTCGGC 242
DB 611 CCCAAGCCCCCAGCGGCTCCAAAGCTCCAGGCGCTCCCGAGCGACCGGTGCTCGGC 670
QY 243 CTTCTCTTGGTTCAGAAAGTCCGCCCTGGGGGAGTTCGTCCTCAAGGGTTCTCGAA 302
DB 671 CTTCTCTTGGTTCAGAAAGTCCGCCCTGGGGGAGTTCGTCCTCAAGGGTTCTCGAA 730
QY 303 AGAATCTGAGGGCGCAGTCTCTGACGAGGGAATCTCTGTGTAGCTTTGGAAGCG 362
DB 731 AGAATCTGAGGGCGCAGTCTCTGACGAGGGAATCTCTGTGTAGCTTTGGAAGCG 790
QY 363 CCAGCCCCCAGAAG 375
DB 791 CCAGCCCCCAGAAG 803

RESULT 2
AL534136 1078 bp mRNA linear EST 24-MAR-2004
LOCUS CSODP005YI02 5-PRIME, mRNA sequence.
DEFINITION AL534136
ACCESSION AL534136 GI:45709952
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1078)
Li, W.-B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 13, 2001 this sequence version replaced gi:30539643.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 6147.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CSODP005B01QP1&c=6147.r.
Location/Qualifiers

FEATURES

BP202832 582 bp mRNA linear EST 14-SEP-2004
BP202832 Sugano cDNA library, amygdala Homo sapiens cDNA clone
AMR09622, mRNA sequence.
ACCESSION BP202832
VERSION BP202832.1 GI:52054220
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 582)
Suzuki, Y., Yamashita, R., Shiota, M., Sakakibara, Y., Chiba, J.,
Mizushima-Yamamoto, J., Nakai, K. and Sugano, S.
Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
Genome Res. 14 (9), 1711-1718 (2004)
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.
Location/Qualifiers

1. 582
source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="AMR09622"
/tissue_type="amygdala"
/clone_lib="Sugano cDNA library, amygdala"

ORIGIN

Query Match 74.5%; Score 279.4; DB 5; Length 582;
Best Local Similarity 99.6%; Pred. No. 7e-58; Indels 0; Gaps 0;
Matches 280; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 95 GTCCTGGTCTGAGCGCGGAGAGCGGCGGAGCGGCTGATCGGCTCCCTCGAAC 154
DB 1 GTCCTGGTCTGAGCGCGGAGAGCGGCGGAGCGGCTGATCGGCGGCTCGAAC 60
QY 155 TGGGAGGTCAGTGGGTCGCTTAGGGCCCAAGCCCCCAGCCCGTCCAAAAGCTCCC 214
DB 61 TGGGAGGTCAGTGGGTCGCTTAGGGCCCAAGCCCCCAGCCCGTCCAAAAGCTCCC 120
QY 215 AGGGCTCCCGAGCGCGGTCGCGCCCTCTCTCGGTGAGAAAGTCCGCCCTGGGG 274
DB 121 AGGGCTCCCGAGCGCGGTCGCGCCCTCTCTCGGTGAGAAAGTCCGCCCTGGGG 180
QY 275 GCAGTTGTCCTCCAAAGGTTTCTCGAAAGAAATCTGAGAGGCGCGAGTCTTGACCGAGG 334
DB 181 GCAGTTGTCCTCCAAAGGTTTCTCGAAAGAAATCTGAGAGGCGCGAGTCTTGACCGAGG 240
QY 335 GAATCTCTGTGTAGCTTGGAGCGCGCCAGCCCGCAGAG 375
DB 241 GAATCTCTGTGTAGCTTGGAGCGCGCCAGCCCGCAGAG 281

RESULT 4

BQ066875
LOCUS BQ066875
DEFINITION AGENCOURT_6818578 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5762608
5', mRNA sequence.
ACCESSION BQ066875
VERSION BQ066875.1 GI:19895921
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 972)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-t@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL2813 row: b column: 17
High quality sequence start: 28
High quality sequence stop: 456.
Location/Qualifiers
1. .972
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5762608"
/lab_host="DH10B"
/clone_lib="NIH_MGC 114"
/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains, age range 23-27 yo. Library is oligo-dt
primed and directionally cloned (EcoRV site is destroyed

FEATURES

source
1. .935
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR14N09"

ORIGIN

Query Match 57.0%; Score 213.8; DB 5; Length 972;
Best Local Similarity 90.7%; Pred. No. 8e-42;
Matches 262; Conservative 0; Mismatches 22; Indels 5; Gaps 3;
QY 3 CTCCTTCCGAGCTGAGCTTACCTCGGCGCAAAACGAGCGAGCGAGGGCGCGAGTGAA 62
DB 412 CTCCTTCCGAGCTGAGCTTACCTCGGCGCAAAACGAGCGAGCGAGGGCGCGAGTGAA 471
QY 63 GCTGGAGTTCGGGCTGGGCGGAGCGGCTGCTCGTGGTGTGAGCGCGCGAGAGC 122
DB 472 GCTGGAGTTCGGGCTGGGCGGAGCGGCTGCTCGTGGTGTGAGCGCGCGAGAGC 531
QY 123 GGGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGCGAGGTCAGTGGGTCGCTTAGGG 182
DB 532 GGGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGCGAGGGCCCGTGGGTCGCTTAAGG 591
QY 183 CCAGAGCCCCCAGCCCGGCTCCAAAAGTCCCA-GGGCTTCCCCAGGACCGGTG--CTC 239
DB 592 GCCAAAGCCCCCAGCCCGGCTCCAAAAGTCCCAAGCGCCCTCCCCAGGACCGGGGCTTC 651
QY 240 GGGCTTCCCTT--CGGTGAGAAAGTCGCCCTGGGCGAGTTCGTCTCC 286
DB 652 GGGCTTCCCTTTCGGGCGCAAAAGTCTCTCCCTCGGGGGCGCGCCCC 700

RESULT 5

CNS006XK/c
LOCUS CNS006XK
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC #
BACR14N09 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL066051
VERSION AL066051.1 GI:4945019
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 935)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Herkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osogawa and
Aaron Mammoser in Pister de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
Location/Qualifiers
1. .935
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR14N09"

```
/clone_lib="RPC1-98"
/notes="end : T7"

ORIGIN
Query Match      15.7%; Score 59; DB 9; Length 935;
Best Local Similarity 31.6%; Pred. No. 0.0058;
Matches 107; Conservative 85; Mismatches 145; Indels 2; Gaps 1;

QY 1 CCTCTCTTCAGCTGAGCTTACCTGGCGCAACAGCAGAGGAGGCGCGAGTGG 60
Db 864 CCGGSSCCGCGCGSSGGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 805
QY 61 AGCTGAGTTCGGGGTGGGAGGAGCGACTGTTCGTGTGTGTAGCGCGCGGAGA 120
Db 804 GCGCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 745
QY 121 GCGGCGCGCGCGCGCGCTGATCGCTCCCTCGCAACTGGGAGGTCCAGTGGGCTCTTAG 180
Db 744 SCGSGGSGCGSCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 685
QY 181 GCGCCAAAGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 238
Db 684 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 625
QY 239 CGGCGCTTCTTCGTCAGAAAGTCCCGCGCGCGCGCGCGCGCGCGCGCG 298
Db 624 CCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 565
QY 299 CGAAAGATCTCAGAGGCGCGAGTCTTTCGACCGAGGGA 337
Db 564 MCRAGSGWGSAGSGRCGAGCGSGSGSGSGSGSGSGSGSGSGSGSGSG 526

RESULT 6
CNS00397/c
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence T7 end of BAC #
BACK08A10 of RPC1-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION
AL063912
VERSION
KEYWORDS
SOURCE
Drosophila melanogaster (fruit fly)
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 1101)
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammosser in Peter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPC1-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila\_bac.htm.
Location/Qualifiers
1. .1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"

FEATURES
source
/clone_lib="RPC1-98"
/notes="end : T7"

/clone="BACR08A10"
/clone_lib="RPC1-98"
/notes="end : T7"

ORIGIN
Query Match      14.3%; Score 53.8; DB 9; Length 1101;
Best Local Similarity 45.7%; Pred. No. 0.011; 84; Indels 0; Gaps 0;
Matches 91; Conservative 24; Mismatches 84; Indels 0; Gaps 0;

QY 29 GCGCAAAACGACGAGGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 88
Db 849 GGGGGGRRGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 790
QY 89 GCGACTGTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 148
Db 789 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 730
QY 149 TCGAACTCGGGAGGTCCAGTGGGGTTCGTTAGGGCCCAAGCCCGGCTCCAAA 208
Db 729 AGGACRSGSGVCGGGSGGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 670
QY 209 GCTCCAGGCGCTCCCGCAG 227
Db 669 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 651

RESULT 7
AG381986
LOCUS
DEFINITION
Mus musculus molossinus DNA, clone:MSMg01-192D21.TJ, genomic survey
sequence.
ACCESSION
AG381986
VERSION
AG381986.1 GI:47993191
KEYWORDS
SOURCE
Mus musculus molossinus
ORGANISM
Mus musculus molossinus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1
Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
BAC end Sequences of Library MSMg01
Unpublished
REFERENCE
2 (bases 1 to 2243)
Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
Direct Submission
Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsuri-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the mouse BAC library MSMg01. For BAC
library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
Tsukuba Institute, Bio Resource Center,
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Kovadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp
PRIMERS
Sequencing : TJ
LIBRARY
Vector : pBACe3.6
R.site 1 : EcoRI
R.site 2 : EcoRI.
Location/Qualifiers
1. .2243
/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57486"
/clone="MSMg01-192D21.TJ"
/sex="male"
/tissue type="mixture of kidney and spleen"
/clone_lib="MSMg01 Mouse Male BAC Library"

FEATURES
source
/clone_lib="RPC1-98"
/notes="end : T7"
```



```
ORIGIN
Query Match      14.2%; Score 53.2; DB 9; Length 2243;
Best Local Similarity 51.0%; Pred. No. 0.016;
Matches 100; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 24 CCTGTGGCGCAACGACGAGCGAGGGCGCGAGTGGAACTGGAGTTCCGGGGTGGCG 83
DB 1905 CGCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1964
QY 84 GCGAGGCGACTGTCCTGCTGTAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 143
DB 1965 CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2024
QY 144 CTCCTCTGAACCTGGGAGGTCCAGTGGGGTCCGTTAGGGGCCCAAGCCCGCGCTC 203
DB 2025 CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2084
QY 204 CAAAGCTCCAGGGC 219
DB 2085 CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG

RESULT 8
AG060136/c
LOCUS
DEFINITION Pan troglodytes DNA, clone: PTB-047116.F, genomic survey sequence.
ACCESSION AG060136
VERSION AG060136.1 GI:16611366
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1
AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE BAC end sequences of Library PTB
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 693)
AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
DIRECT SUBMISSION
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-cho,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
E-mail:chimpses@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: -21M13
LIBRARY
Vector : PKS145
R.Site 1 : SacI
R.Site 2 : SacI.
FEATURES
source
Location/Qualifiers
1. .693
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-047116.F"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"

ORIGIN
Query Match      14.1%; Score 53; DB 9; Length 693;
Best Local Similarity 49.4%; Pred. No. 0.017;
Matches 131; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY 28 GGGCGCAACGACGAGCGAGGGCGCGAGTGGAACTGGAGTTCCGGGGTGGCGGGA 87

ORIGIN
Query Match      13.8%; Score 51.8; DB 9; Length 935;
Best Local Similarity 33.2%; Pred. No. 0.033;
Matches 86; Conservative 58; Mismatches 115; Indels 0; Gaps 0;

QY 28 GGGCGCAACGAGCGAGGGCGCGAGTGGAACTGGAGTTCCGGGGTGGCGGGA 87
DB 669 GGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 728
```

```
Db 647 GGGGCGGGCGCGCGCGCGGGTTTGGGNGGCGGCGCGGTGCGCGGGGGGGGCG 588
QY 88 GGCACACTGTCGTGTGCTAGCGCGCGGAGAGCGGCGCGAGCGGTGATCGGTCC 147
DB 587 GGGGGGGGGCGCGGGGGGGCGCGCGCGGNNCTGCGCGCGCGCGCGCGCGCGCG 528
QY 148 CTCGAACCTGGGAGGTCCAGTGGGGTCCAGTGGGGCCAAAGCCCCCACCCTCCAAA 207
DB 527 GGGTCGGTGGGGGGCGCGGGGGCGGGGGGGGGTGGCGGGGGGGGGGGCGCGCG 468
QY 208 AGCTCCACAGGCGCTCCCGAGGCAACCGTGTCTCGGCCCTTCTTTCGTCGAGAAAGTCGCC 267
DB 467 TGGGGGGGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 408
QY 268 CTGGGGGGCAGTTCGTCCCAAGGG 292
DB 407 GCGGGGGGGCGTGGGGGGCGCGCGGG 383

CNS006XK 935 bp DNA linear GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence T7 end of BAC #
BACr14N09 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL066051
VERSION AL066051.1 GI:4945019
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 935)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammoler in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain v2; cn bw ap, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES
source
Location/Qualifiers
1. .935
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACr14N09"
/clone_lib="RPCI-98"
/note="end : T7"

ORIGIN
Query Match      13.8%; Score 51.8; DB 9; Length 935;
Best Local Similarity 33.2%; Pred. No. 0.033;
Matches 86; Conservative 58; Mismatches 115; Indels 0; Gaps 0;

QY 28 GGGCGCAACGAGCGAGGGCGCGAGTGGAACTGGAGTTCCGGGGTGGCGGGA 87
DB 669 GGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 728
```



```
Db 692 TGSTSSSSSSSTSSSSVSSGSKSTBSGSSBSSSSSSSSSTSSBBSCTSSSSSS 751
QY 185 CAAGACCCACCCAGCGCTCCAAAAGCTCCAGGCGCTCCCCAGGACCGGTGTCGCGCC 244
Db 752 SYSSSTCCCTCCCSYSSSTSSSSSTSSGSGSSSSVGTSSSDSTSTCCSCCCYM 811
QY 245 TTCCTTCGGTCAGAAAAGTCGCCCCCTCGGGGCGAGTTCGTCCTCCAAAGGTTTCTCGAAAG 304
Db 812 CTCSTYBMCYTSSTCGSSSSSGKGVTKCGCGGSSSTGMBGTSSACSSSSSSCS 871
QY 305 AATCTGAGAGGGCGCAGTCTTGACCGAGGGAATC 339
Db 872 SSSVSSSSKSSASSSSVSSSGSGSGVSSNSSSASKS 906

RESULT 12
LOCUS CL493826
DEFINITION SAIL_586_F04.v1 SAIL Collection Arabidopsis thaliana genomic clone
ACCESSION SAIL_586_F04.v1, genomic survey sequence.
VERSION CL493826
KEYWORDS CL493826.1 GI:45982408
SOURCE GSS.
ORGANISM Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 890)
AUTHORS Sessions,A., Burke,E., Presting,G., Aux,G., McElver,J., Patton,D.,
Dietrich,B., Ho,P., Bacwaden,J., Ko,C., Clarke,J.D., Cotton,D.,
Bullis,D., Snell,J., Miguel,T., Hutchison,D., Kimmerly,B.,
Mitzel,T., Katagiri,P., Glazebrook,J., Law,M. and Goff,S.A.
A high-throughput Arabidopsis reverse genetics system
Plant Cell 14 (12), 2985-2994 (2002)
22356987
12468722
COMMENT Contact: Sessions A
Applied Trait Genetics
Syngenta Biotechnology Inc.
3054 Cornwallis Rd., Research Triangle Park, NC 27709, USA
Email: allen.sessions@syngenta.com
ABRC Stock Number CS824955; T-DNA left border flanking sequences of
Syngenta Arabidopsis Insertion Library (SAIL) lines are available
through the Arabidopsis Biological Resource Center (ABRC).
Sequences represent a pool of amplified genomic regions and not
single contiguous sequences.
Class: T-DNA tagged.
FEATURES
source Location/Qualifiers
1. .890
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/ecotype="Columbia"
/db_xref="taxon:3702"
/clone="SAIL_586_F04.v1"
/clone_lib="SAIL Collection"
/notes="T-DNA left border sequences were isolated using a
modified TAIL-PCR strategy"
ORIGIN
Query Match 13.4%; Score 50.4; DB 9; Length 890;
Best Local Similarity 48.6%; Pred. No. 0.073;
Matches 126; Conservative 0; Mismatches 133; Indels 0; Gaps 0;
QY 28 GGCGGCAACAGCAGGCGAGCGCGCGAGTGGAGCTGGAGTTCGGGGTGGCGGGGA 87
Db 615 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 674
QY 88 GGGGACTGTCTCGTGGTCTCAGCGCGCGCGAGAGCGGCGGAGCGGCTGATCGGCTCC 147
Db 675 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 734
QY 148 CTCGAACTGGGGAGGTTCAGTGGGGTTCGCTTAGGGGCCCAAGACCCGCCCGGCTCCAA 207
```

```
Db 735 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 794
QY 208 AGCTCCAGGCGCTCCCGGACCGGTGTCGCGCCCTCTCTCGGTTCAGAAAGTCGCCC 267
Db 795 CCCCCGGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 854
QY 268 CTTGGGGGCGAGTTCGTCCC 286
Db 855 CGCGGGCGCGCGCGCGCG 873

RESULT 13
LOCUS AG127812
DEFINITION Pan troglodytes DNA, clone: PTB-138M14.R, genomic survey sequence.
ACCESSION AG127812
VERSION AG127812.1 GI:16656977
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
REFERENCE 1
AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
BAC end sequences of Library PTB
Unpublished
2 (bases 1 to 964)
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
[E-mail:chimps@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170]
Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: M13Rev
LIBRARY Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI,
FEATURES
source Location/Qualifiers
1. .964
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-138M14.R"
/cell="male"
/sex="male"
/clone_lib="PTB Chimpanzee Male BAC Library"
ORIGIN
Query Match 13.3%; Score 50; DB 9; Length 964;
Best Local Similarity 48.1%; Pred. No. 0.092;
Matches 116; Conservative 0; Mismatches 125; Indels 0; Gaps 0;
QY 38 GAGCGAGGCGGGCGCGAGTGGAGTCCGGGTTCGGGGTGGCGGAGCGGACTGTC 97
Db 340 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 399
QY 98 CFTGTGTCAGCGCGCGGAGAGCGGCGCGAGCGGTGATCGGCTCCCTCGAACTGG 157
Db 400 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 459
QY 158 GGNAGTTCAGTGGGTTCGCTTAGGGGCCCAAGACCCGCCCGGCTCCAAAGTCCAGG 217
Db 460 GNGGGGNGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 519
```


Search completed: April 7, 2005, 14:55:34
Job time : 2245.21 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2005, 07:44:40 ; Search time 3090.98 Seconds
(without alignments)
10158.276 Million cell updates/sec

Title: US-09-977-579-4_COPY_376_1023

Perfect score: 648

Sequence: 1 atgcctgccttcattagatt.....cggtagcagtgaggagaatag 648

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|--------------|--------------------|
| 1 | 648 | 100.0 | 1261 | 6 AR359850 | AR359850 Sequence |
| 2 | 648 | 100.0 | 1261 | 6 AX039100 | AX039100 Sequence |
| 3 | 648 | 100.0 | 1261 | 9 HSA243396 | AJ243396 Homo sapi |
| 4 | 646.4 | 99.8 | 5306 | 9 AB032984 | AB032984 Homo sapi |
| 5 | 627.2 | 96.8 | 3296 | 9 AB097521 | AB097521 Macaca fa |
| 6 | 592.4 | 91.4 | 606 | 6 CQ728741 | CQ728741 Sequence |
| 7 | 537.2 | 82.9 | 670 | 10 AY049036 | AY049036 Mus muscu |
| 8 | 536.6 | 82.8 | 2220 | 6 AR359849 | AR359849 Sequence |
| 9 | 536.6 | 82.8 | 2220 | 6 AX039099 | AX039099 Sequence |
| 10 | 536.6 | 82.8 | 2220 | 10 RN0243395 | AJ243395 Rattus no |
| 11 | 536.6 | 82.8 | 2632 | 6 AX048005 | AX048005 Sequence |
| 12 | 536.6 | 82.8 | 3107 | 10 AF378093 | AF378093 Rattus no |
| 13 | 536.6 | 82.8 | 3108 | 6 AX047984 | AX047984 Sequence |
| 14 | 536.6 | 82.8 | 4025 | 10 AK173115 | AK173115 Mus muscu |
| 15 | 536.6 | 82.8 | 4169 | 10 BC053919 | BC053919 Mus muscu |
| 16 | 536.6 | 82.8 | 4176 | 10 BC058636 | BC058636 Mus muscu |
| 17 | 534.6 | 82.5 | 645 | 6 AX048004 | AX048004 Sequence |
| 18 | 533.4 | 82.3 | 3910 | 10 BC070899 | BC070899 Rattus no |
| 19 | 422 | 65.1 | 471 | 6 BD059018 | BD059018 Secreted |

| | | | | | | |
|----|-------|------|--------|----|-----------|-------------|
| 20 | 368 | 56.8 | 922 | 5 | BX935227 | Gallus ga |
| 21 | 332.2 | 51.3 | 912 | 5 | BX931708 | Gallus ga |
| 22 | 264.6 | 40.8 | 1654 | 5 | BC077295 | Xenopus l |
| 23 | 229.6 | 35.4 | 127347 | 2 | AF000682 | Homo sapi |
| 24 | 229.6 | 35.4 | 172546 | 2 | AC024604 | Homo sapi |
| 25 | 229.6 | 35.4 | 178169 | 9 | AP002765 | Homo sapi |
| 26 | 229.6 | 35.4 | 178431 | 2 | AP002749 | Homo sapi |
| 27 | 229.6 | 35.4 | 181471 | 9 | AC069539 | Homo sapi |
| 28 | 228 | 35.2 | 144833 | 2 | AC063921 | Homo sapi |
| 29 | 228 | 35.2 | 149800 | 2 | AC021981 | Homo sapi |
| 30 | 204.2 | 31.5 | 174285 | 2 | AC145484 | Lemur cat |
| 31 | 179 | 27.6 | 140307 | 2 | AC135353 | Mus muscu |
| 32 | 179 | 27.6 | 238861 | 2 | AC148331 | Mus muscu |
| 33 | 174.2 | 26.9 | 214306 | 2 | AC128723 | Rattus no |
| 34 | 167.8 | 25.9 | 144833 | 2 | AC063921 | Homo sapi |
| 35 | 148.2 | 22.9 | 657 | 4 | OCU35382 | Oryctolagus |
| 36 | 148.2 | 22.9 | 657 | 6 | AX048006 | Sequence |
| 37 | 145.4 | 22.4 | 1404 | 9 | HUMSCN1BA | Human sodiu |
| 38 | 143.8 | 22.2 | 617 | 6 | CQ722293 | Sequence |
| 39 | 143.8 | 22.2 | 1335 | 9 | HUMVGS1B | Sequence |
| 40 | 143.8 | 22.2 | 1525 | 9 | BC067122 | Homo sapi |
| 41 | 138.6 | 21.4 | 1125 | 10 | MMU46681 | Mus musculu |
| 42 | 138.6 | 21.4 | 1340 | 10 | MMU85786 | Mus musculu |
| 43 | 138.6 | 21.4 | 1490 | 6 | AX401978 | Sequence |
| 44 | 138.6 | 21.4 | 1490 | 10 | RATSCHB1B | Rattus norv |
| 45 | 138.6 | 21.4 | 1557 | 10 | BC039140 | Mus muscu |

ALIGNMENTS

| | | | | | | |
|------------|---|------------------------------------|---------|-----|--------|-----------------|
| RESULT 1 | AR359850 | Sequence 4 from patent US 6593565. | 1261 bp | DNA | linear | PAT 17-AUG-2003 |
| LOCUS | AR359850 | Sequence 4 from patent US 6593565. | 1261 bp | DNA | linear | PAT 17-AUG-2003 |
| DEFINITION | AR359850 | Sequence 4 from patent US 6593565. | 1261 bp | DNA | linear | PAT 17-AUG-2003 |
| ACCESSION | AR359850 | Sequence 4 from patent US 6593565. | 1261 bp | DNA | linear | PAT 17-AUG-2003 |
| VERSION | AR359850.1 | GI:33766660 | 1261 bp | DNA | linear | PAT 17-AUG-2003 |
| KEYWORDS | Unknown. | | | | | |
| SOURCE | Unknown. | | | | | |
| ORGANISM | Unknown. | | | | | |
| REFERENCE | 1 (bases 1 to 1261) | | | | | |
| AUTHORS | Heslin, P. and Lynam, N.R. | | | | | |
| TITLE | Vehicle interior rearview mirror assembly including an accessory-containing housing | | | | | |
| JOURNAL | Patent: US 6593565-A 4 15-JUL-2003; | | | | | |
| FEATURES | Location/Qualifiers | | | | | |
| source | 1. .1261 | | | | | |
| organism | /organism="unknown" | | | | | |
| mol_type | /mol_type="genomic DNA" | | | | | |

ORIGIN

| | | | | |
|-----------------------|-----------------|---|-----------|--------------|
| Query Match | 100.0%; | Score 648; | DB 6; | Length 1261; |
| Best Local Similarity | 100.0%; | Pred. No. 2.1e-142; | | |
| Matches 648; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |
| Qy | 1 | ATGCCTGCCTTCAATAGATTGTTCCCTCGCTGCTTCTCTCGTGGCTTCTCTACTGGTCAGT | 60 | |
| Db | 376 | ATGCCTGCCTTCAATAGATTGTTCCCTCGCTGCTTCTCTCGTGGCTTCTCTACTGGTCAGT | 435 | |
| Qy | 61 | GTCTGCTTCCCTGTGTGTGTGGAAGTCCCTCGGAGACGGAGCGCGTCGAGGGCAACCCC | 120 | |
| Db | 436 | GTCTGCTTCCCTGTGTGTGTGGAAGTCCCTCGGAGACGGAGCGCGTCGAGGGCAACCCC | 495 | |
| Qy | 121 | ATGAAGTCGGCTGCATCTCTCGATGAAGAGAGAGAGGTGGAGGCCACACCGTGGTG | 180 | |
| Db | 496 | ATGAAGTCGGCTGCATCTCTCGATGAAGAGAGAGAGGTGGAGGCCACACCGTGGTG | 555 | |
| Qy | 181 | GAATGGTTCTACAGGCGCGGAGGCGGTAAAGATTTCCTTATTTACGAGTATCGGAATGGC | 240 | |
| Db | 556 | GAATGGTTCTACAGGCGCGGAGGCGGTAAAGATTTCCTTATTTACGAGTATCGGAATGGC | 615 | |
| Qy | 241 | CACCAGGAGGTGGAGAGACCCCTTTTCAGGGCGGCTTCGATGGAATGCGACGAGGACCTG | 300 | |

```
Db 616 CACCAGGAGTGGAGAGCCCTTTTCAGGGCGCCTCAGTGGATGGCAGAGGACCTG 675
Qy 301 CAGGACGTGTCCATCATCTGTGCTCAACGTCACTCTGAACGACTCTGGCCCTCTACACCTGC 360
Db 676 CAGGACGTGTCCATCATCTGTGCTCAACGTCACTCTGAACGACTCTGGCCCTCTACACCTGC 735
Qy 361 AATGTGTCCTCCGGAGTTTGAAGTTGAGGCGCATCGGCCCTTTGTGAAGACGACGCGGCTG 420
Db 736 AATGTGTCCTCCGGAGTTTGAAGTTGAGGCGCATCGGCCCTTTGTGAAGACGACGCGGCTG 795
Qy 421 ATCCCCCTAAGAGTCAACGAGAGGCTGGAGAGGACTTCACCTCTGTGCTCTCAGAAATC 480
Db 796 ATCCCCCTAAGAGTCAACGAGAGGCTGGAGAGGACTTCACCTCTGTGCTCTCAGAAATC 855
Qy 481 ATGATGTACATCCTTCTGTGCTTCTCCTCACCCTGTGGCTGCTCATCGAGATGATATATTC 540
Db 856 ATGATGTACATCCTTCTGTGCTTCTCCTCACCCTGTGGCTGCTCATCGAGATGATATATTC 915
Qy 541 TACAGAAAGGTCTCAAAAGCCGAGAGGAGGCCCAAGAAACGCGTCTGACTACCTTGCC 600
Db 916 TACAGAAAGGTCTCAAAAGCCGAGAGGAGGCCCAAGAAACGCGTCTGACTACCTTGCC 975
Qy 601 ATCCCATCTGAGAACAGAGAGACTCTCGGTTACCAAGTGGAGGAATAG 648
Db 976 ATCCCATCTGAGAACAGAGAGACTCTCGGTTACCAAGTGGAGGAATAG 1023

RESULT 2
AX039100
LOCUS AX039100 1261 bp DNA linear PAT 18-NOV-2000
DEFINITION Sequence 4 from Patent WO0063367.
ACCESSION AX039100
VERSION AX039100.1 GI:11229276
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
TITLE Cox, P., Dixon, A., Jackson, A. and Morgan, K.
JOURNAL A novel family of beta sub-unit proteins from a voltage-gated sodi
diagnostic uses there of
Patent: WO 0063367-A 4 26-OCT-2000;
WARNER-LAMBERT COMPANY (US) ; Cambridge University Technical
SERVICES Limited (GB)
FEATURES
Location/Qualifiers
source 1..1261
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 648; DB 6; Length 1261;
Best Local Similarity 100.0%; Pred. No. 2.1e-142; Indels 0; Gaps 0;
Matches 648; Conservative 0; Mismatches 0;

Qy 1 ATGCGTCCTTCAATAGATTGTTTCCCTGGCTTCTCGTCTTACTTACTTGGGTCACT 60
Db 376 ATGCGTCCTTCAATAGATTGTTTCCCTGGCTTCTCGTCTTACTTACTTGGGTCACT 435
Qy 61 GTCTGCTTCCCTGTGTGTAAGTGCCTCGGAGACGAGGCGCGTGGAGGCAACCC 120
Db 436 GTCTGCTTCCCTGTGTGTAAGTGCCTCGGAGACGAGGCGCGTGGAGGCAACCC 495
Qy 121 ATGAAGCTCGCTGCATCTCTCGATGAGAGAGGAGTGGAGCCACCGTGGTG 180
Db 496 ATGAAGCTCGCTGCATCTCTCGATGAGAGAGGAGTGGAGCCACCGTGGTG 555
Qy 181 GAATGTTCTACAGGCCCGGAGGCGGTAAAGATTTCTTTATTTACGAGTATCGGAATGC 240
Db 556 GAATGTTCTACAGGCCCGGAGGCGGTAAAGATTTCTTTATTTACGAGTATCGGAATGC 615
```

```
Qy 241 CACCAGGAGTGGAGAGCCCTTTTCAGGGCGCCTCAGTGGATGGCAGAGGACCTG 300
Db 616 CACCAGGAGTGGAGAGCCCTTTTCAGGGCGCCTCAGTGGATGGCAGAGGACCTG 675
Qy 301 CAGGACGTGTCCATCATCTGTGCTCAACGTCACTCTGAACGACTCTGGCCCTCTACACCTGC 360
Db 676 CAGGACGTGTCCATCATCTGTGCTCAACGTCACTCTGAACGACTCTGGCCCTCTACACCTGC 735
Qy 361 AATGTGTCCTCCGGAGTTTGAAGTTGAGGCGCATCGGCCCTTTGTGAAGACGACGCGGCTG 420
Db 736 AATGTGTCCTCCGGAGTTTGAAGTTGAGGCGCATCGGCCCTTTGTGAAGACGACGCGGCTG 795
Qy 421 ATCCCCCTAAGAGTCAACGAGAGGCTGGAGAGGACTTCACCTCTGTGCTCTCAGAAATC 480
Db 796 ATCCCCCTAAGAGTCAACGAGAGGCTGGAGAGGACTTCACCTCTGTGCTCTCAGAAATC 855
Qy 481 ATGATGTACATCCTTCTGTGCTTCTCCTCACCCTGTGGCTGCTCATCGAGATGATATATTC 540
Db 856 ATGATGTACATCCTTCTGTGCTTCTCCTCACCCTGTGGCTGCTCATCGAGATGATATATTC 915
Qy 541 TACAGAAAGGTCTCAAAAGCCGAGAGGAGGCCCAAGAAACGCGTCTGACTACCTTGCC 600
Db 916 TACAGAAAGGTCTCAAAAGCCGAGAGGAGGCCCAAGAAACGCGTCTGACTACCTTGCC 975
Qy 601 ATCCCATCTGAGAACAGAGAGACTCTCGGTTACCAAGTGGAGGAATAG 648
Db 976 ATCCCATCTGAGAACAGAGAGACTCTCGGTTACCAAGTGGAGGAATAG 1023
```

RESULT 3
HSA243396
LOCUS
DEFINITION Homo sapiens mRNA for voltage-gated sodium channel beta-3 subunit (scn3b gene).

ACCESSION AJ243396
VERSION AJ243396.2 GI:7242612
KEYWORDS scn3b gene; voltage-gated sodium channel beta-3 subunit.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Morgan, K., Stevens, E.B., Shaw, B., Cox, P., Dixon, A.K., Lee, K., Finnock, R.D., Hynes, J., Richardson, P.J., Mizuguchi, K. and Jackson, A.P.

TITLE beta 3: an additional auxiliary subunit of the voltage-sensitive sodium channel that modulates channel gating with distinct kinetics
Proc. Natl. Acad. Sci. U.S.A. 97 (5), 2308-2313 (2000)

JOURNAL
MEDLINE 20160948
PubMed 10688874

REFERENCE
Morgan, K.
Direct Submission

JOURNAL
Submitted (28-JUN-1999) Morgan K., Biochemistry, University of Cambridge, Tennis Court Road, Cambridge, CB2 1QW, UNITED KINGDOM
Revised by [4]

REMARK 3 (bases 1 to 1261)

AUTHORS
Morgan, K.
Direct Submission

JOURNAL
Submitted (13-MAR-2000) Morgan K., Biochemistry, University of Cambridge, Tennis Court Road, Cambridge, CB2 1QW, UNITED KINGDOM
On Mar 14, 2000 this sequence version replaced gi:7160974.

COMMENT
Location/Qualifiers

source
1..1261
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="brain"
1..1261
/gene="scn3b"
-1..375
/gene="scn3b"


```

CDS
376..1024
/gene="scn3b"
/function="modulates channel gating kinetics"
/codon_start=1
/evidence=experimental
/product="voltage-gated sodium channel beta-3 subunit"
/protein_id="CAB76825.1"
/db_xref="GI:7160975"
/db_xref="GOA:Q9NV72"
/db_xref="UniProt/Swiss-Prot:Q9NV72"
/translation="MPAPNRLFPPLASLVLIYVSVCPVCEVPSSETEAVQGNPKMLR
CSMKRKEEVATTVWFYRPEGGKDFLIYRNGHVEVSPFQGRQWNGSKDLOD
VSIIVLVNLTLDLSGLYTCNVSREFEFAHRPFVKTTRLIPLRVTEAGEDFTSVSSI
MMYILVFLTLMLLIEMIYCYRKVSKAEAAQENASDYLAIIPSENKENSAPVVEE"
1025..>1261
/gene="scn3b"

3'UTR
ORIGIN
Query Match 100.0%; Score 648; DB 9; Length 1261;
Best Local Similarity 100.0%; Pred. No. 2.1e-142;
Matches 648; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCTGCTTCAATAGATTGTTTCCCTGGCTTCTCTCGTCTATCTACTGGTCACT 60
Db |||||
Qy 376 ATGCTGCTTCAATAGATTGTTTCCCTGGCTTCTCTCGTCTATCTACTGGTCACT 435
Db |||||
Qy 61 GTCTGCTTCCCTGTGTGGAAGTGCCTCGGAGACGGAGCGCGTGCAGGGCAACCCC 120
Db |||||
Qy 436 GTCTGCTTCCCTGTGTGGAAGTGCCTCGGAGACGGAGCGCGTGCAGGGCAACCCC 495
Db |||||
Qy 121 ATGAAGCTGCGTCATCTCTCGATGAAGAGAGAGAGGTGGAGGCCACCCACGGTGGT 180
Db |||||
Qy 496 ATGAAGCTGCGTCATCTCTCGATGAAGAGAGAGAGGTGGAGGCCACCCACGGTGGT 555
Db |||||
Qy 181 GAATGGTTTACAGCCCGAGGGCGGTAAAGATTCTTATTTACGAGTATCGGAATGGC 240
Db |||||
Qy 556 GAATGGTTTACAGCCCGAGGGCGGTAAAGATTCTTATTTACGAGTATCGGAATGGC 615
Db |||||
Qy 241 CACCAGAGGTGGAGAGCCCTTTACGGGGCGCTCGAGTGAATGTCAGCAAGAGCCTG 300
Db |||||
Qy 616 CACCAGAGGTGGAGAGCCCTTTACGGGGCGCTCGAGTGAATGTCAGCAAGAGCCTG 675
Db |||||
Qy 301 CAGGAGGTGCTCACTGCTCAACGTCCTCAACGTCCTCAACGTCCTCAACGTCCT 360
Db |||||
Qy 676 CAGGAGGTGCTCACTGCTCAACGTCCTCAACGTCCTCAACGTCCTCAACGTCCT 735
Db |||||
Qy 361 AATGTGTCGGGAGTTGATTTGAGCGGCATCGGCCCTTTGTGAAGACAGCGCGCTG 420
Db |||||
Qy 736 AATGTGTCGGGAGTTGATTTGAGCGGCATCGGCCCTTTGTGAAGACAGCGCGCTG 795
Db |||||
Qy 421 ATCCCTTAAGAGTCAAGAGGAGGCTGGAGAGCTTCACTCTGTGTTCTCAGAAATC 480
Db |||||
Qy 796 ATCCCTTAAGAGTCAAGAGGAGGCTGGAGAGCTTCACTCTGTGTTCTCAGAAATC 855
Db |||||
Qy 481 ATGATGTACATCTTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 540
Db |||||
Qy 856 ATGATGTACATCTTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 915
Db |||||
Qy 541 TACAGAAAGTCTCAAAAGCGGAGAGAGCGGCCCAAGAAACCGCTCTGACTACCTTGGC 600
Db |||||
Qy 916 TACAGAAAGTCTCAAAAGCGGAGAGAGCGGCCCAAGAAACCGCTCTGACTACCTTGGC 975
Db |||||
Qy 601 ATCCCATCTGAGACACAGGAGACTCTCGGTACCGTACCGTGGAGGAATAG 648
Db |||||
Qy 976 ATCCCATCTGAGACACAGGAGACTCTCGGTACCGTGGAGGAATAG 1023
Db |||||

RESULT 4
AB032984
LOCUS 5306 bp mRNA linear PRI 11-NOV-1999
DEFINITION Homo sapiens mRNA for KIAA1158 protein, partial cds.
ACCESSION AB032984
VERSION AB032984.1 GI:6330135

```

```

KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Hiroseawa.M., Nagase.T., Ishikawa.K., Kikuno.R., Nomura.N. and
Ohara.O.
TITLE Characterization of cDNA clones selected by the GeneMark analysis
from size-fractionated cDNA libraries from human brain
JOURNAL DNA Res. 6 (5), 329-336 (1999)
MEDLINE 20039618
PUBMED 10574461
REFERENCE 2 (bases 1 to 5306)
AUTHORS Ohara.O., Nagase.T. and Kikuno.R.
TITLE Direct Submission
JOURNAL Submitted (04-OCT-1999) Osamu Ohara, Kazusa DNA Research Institute,
Laboratory of DNA Technology; 1532-3 Yana, Kisarazu, Chiba
292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp,
URL:http://www.kazusa.or.jp/huge/, Tel:+81-438-52-3913,
Fax:+81-438-52-3914)
FEATURES
Location/Qualifiers
1..5306
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="bj00081"
/sex="male"
/tissue_type="brain"
/clone_lib="pBluescriptII SK plus"
/dev_stage="adult"
1..5306
/gene="KIAA1158"
<1..695
/gene="KIAA1158"
/notes="Start codon is not identified."
/codon_start=3
/product="KIAA1158 protein"
/protein_id="BAA86472.1"
/db_xref="GI:6330136"
/translation="RSWGSGSLGSRQPKMPFNRLFPPLASLVLIYVSVCPVCEV
PSTEAVQGNPKMLRCISMKRKEEVATTVWFYRPEGGKDFLIYRNGHVEVSP
FQGRQWNGSKDLODYSITVNLNDSGLYTCNVSREFEFAHRPFVKTTRLIPLRV
TEBAGDFTSVSVEIMWYILLVFLTLMLLIEMIYCYRKVSKAEAAQENASDYLAI
PSENKENSAPVVEE"
ORIGIN
Query Match 99.8%; Score 646.4; DB 9; Length 5306;
Best Local Similarity 99.8%; Pred. No. 4.7e-142;
Matches 647; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGCTGCTTCAATAGATTGTTTCCCTGGCTTCTCTCGTCTATCTACTGGTCACT 60
Db |||||
Qy 48 ATGCTGCTTCAATAGATTGTTTCCCTGGCTTCTCTCGTCTATCTACTGGTCACT 107
Db |||||
Qy 61 GTCTGCTTCCCTGTGTGGAAGTGCCTCGGAGACGGAGCGCGTGCAGGGCAACCCC 120
Db |||||
Qy 108 GTCTGCTTCCCTGTGTGGAAGTGCCTCGGAGACGGAGCGCGTGCAGGGCAACCCC 167
Db |||||
Qy 121 ATGAAGCTGCGTCATCTCTCGATGAAGAGAGAGGTGGAGGCCACCCACGGTGGT 180
Db |||||
Qy 168 ATGAAGCTGCGTCATCTCTCGATGAAGAGAGAGGTGGAGGCCACCCACGGTGGT 227
Db |||||
Qy 181 GAATGGTTTACAGCCCGAGGGCGGTAAAGATTCTTATTTACGAGTATCGGAATGGC 240
Db |||||
Qy 228 GAATGGTTTACAGCCCGAGGGCGGTAAAGATTCTTATTTACGAGTATCGGAATGGC 287
Db |||||
Qy 241 CACCAGAGGTGGAGAGCCCTTTACGGGGCGCTGCAGTGAATGTCAGCAAGAGCCTG 300
Db |||||
Qy 288 CACCAGAGGTGGAGAGCCCTTTACGGGGCGCTGCAGTGAATGTCAGCAAGAGCCTG 347
Db |||||
Qy 301 CAGGAGGTGCTCACTGCTTCAAGCTCACTCTGAACGACTCTGCGCTTACACCTGC 360
Db |||||

```



```

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1
REFERENCE Venter, C.J., Adams M.C., Li, P.W. and Myers, E.W.
AUTHORS Kites, such as nucleic acid arrays, comprising a majority of
TITLE humanexons or transcripts, for detecting expression and other uses
JOURNAL Patent: WO 02068579-A 14675 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES Location/Qualifiers
source
1..606
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 91.4%; Score 592.4; DB 6; Length 606;
Best Local Similarity 99.8%; Pred. No. 2.9e-129;
Matches 593; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 55 GTGAGTCTGCTCCCTGCTGTGTGTAAGTGCCTCGGAGACGAGGCGGTGCAGGGC 114
DB 13 GTGAGTCTGCTCCCTGCTGTGTGTAAGTGCCTCGGAGACGAGGCGGTGCAGGGC 72
QY 115 AACCCCATGAAGTGCCTGCTGCTGTCATGAAGAGAGAGAGGAGTGGAGGCCACACG 174
DB 73 AACCCCATGAAGTGCCTGCTGCTGTCATGAAGAGAGAGAGGAGTGGAGGCCACACG 132
QY 175 GTGAGTGAAGTGTCTACAGGCGCGGAGGCGGTGAAGATTCCTATTACGATATCGG 234
DB 133 GTGAGTGAAGTGTCTACAGGCGCGGAGGCGGTGAAGATTCCTATTACGATATCGG 192
QY 235 AATGGCCACACGAGGTGGAGAGCCCTTTTCAGGGCGCCTGCAAGTGGAGTGGAGCAAG 294
DB 193 AATGGCCACACGAGGTGGAGAGCCCTTTTCAGGGCGCCTGCAAGTGGAGTGGAGCAAG 252
QY 295 GACCTGCAGGACGTGTCATCCTGCTCAAGTCACTCTGCAACGCTCTGCAACGCTCTG 354
DB 253 GACCTGCAGGACGTGTCATCCTGCTCAAGTCACTCTGCAACGCTCTGCAACGCTCTG 312
QY 355 ACCTGCAATGTGTCGGGAGTTGAGTTGAGCGGCATCGGCCCTTTGTGAAGACAGC 414
DB 313 ACCTGCAATGTGTCGGGAGTTGAGTTGAGCGGCATCGGCCCTTTGTGAAGACAGC 372
QY 415 CGGCTGATCCCTTAAGAGTCAACGAGGAGGCTCGAGAGACTTCACCTCTGTGGTCTCA 474
DB 373 CGGCTGATCCCTTAAGAGTCAACGAGGAGGCTCGAGAGACTTCACCTCTGTGGTCTCA 432
QY 475 GAAATCATGATGATACATCTTCTGCTTCTCTCAACCTGTGGTGTCTCATCGAGATGATA 534
DB 433 GAAATCATGATGATACATCTTCTGCTTCTCTCAACCTGTGGTGTCTCATCGAGATGATA 492
QY 535 TATTGCTACGAAGGTCTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 594
DB 493 TATTGCTACGAAGGTCTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 552
QY 595 CTGTCATCCATCTGAGAACAGGAGAACTCTGCGGTACAGTGGAGGAATAG 648
DB 553 CTGTCATCCATCTGAGAACAGGAGAACTCTGCGGTACAGTGGAGGAATAG 606

RESULT 7
AY049036 570 bp mRNA linear ROD 01-OCT-2001
LOCUS Mus musculus brain and heart sodium channel beta 3 subunit mRNA,
DEFINITION complete cds.
ACCESSION AY049036
VERSION AY049036.1 GI:15822811
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 670)
REFERENCE
Chen, C., Avery, C., Kazen-Gillespie, K. and Isom, L.L.
AUTHORS Mouse brain and heart beta 3 sodium channel cDNA
JOURNAL Unpublished
TITLE 2 (bases 1 to 670)
AUTHORS Chen, C., Avery, C., Kazen-Gillespie, K. and Isom, L.L.
JOURNAL Direct Submission
TITLE Submitted (30-JUL-2001) Pharmacology, University of Michigan, 1301
JOURNAL MSRB III, Box 0632, Ann Arbor, MI 48109-0632, USA
FEATURES Location/Qualifiers
source
1..670
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
18..665
/notes="auxiliary protein; cell adhesion molecule"
/codon_start=1
/product="brain and heart sodium channel beta 3 subunit"
/protein_id="AAL07512.1"
/db_xref="GI:15822812"
/translators="MPAFNRLPLASLILVYVVRVCPVCPVSETEAVQNSMKLR
CISCKREVEATTVVWVYRPSGKDFLYEYRNGHOEVESPPQGRLOWNSKDLQD
VSTIVNLVNDLSGLTVCNVSREFEFEHPRFVKTRILIPRVTEAGEFTSVVSEI
MMYILLVFLTLWLFIEIMYCYRKVSKABEAXENASDYLAIPSENKENSVPVVEE"
ORIGIN
Query Match 82.9%; Score 537.2; DB 10; Length 670;
Best Local Similarity 89.3%; Pred. No. 3.1e-116;
Matches 578; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
QY 1 ATGCTGCTCCCTCAATAGATTGTTTCCCTGGCTCTCTCGTGTCTATCTACTCTGGTCA 60
DB 18 ATGCTGCTCCCTCAACAGATTGTTTCCCTAGTCTCTCTAGTGTCTATCTACTCTGGT 77
QY 61 GTCTGCTTCCCTGTGTGTGGAAGTCCCTCGAGACGAGGCGGTGAGGGCAACCC 120
DB 78 GTCTGCTTCCCTGTGTGTGGAAGTACCTCGAGACGAGGCGGTGAGGGCAATTC 137
QY 121 ATGAAGCTGCGTCTCTCTGATGAAGAGAGAGAGGAGTGGAGGCGCACCCAGTGTG 180
DB 138 ATGAAGCTGATGATGATCTCTCTGATGAAGAGAGAGGAGTGGAGGCGCACCTGT 197
QY 181 GAATGTTTACAGGCGCGGCGGTAAGATTTCTTATTTACAGATTCGGAATGGC 240
DB 198 GAGTGGTCTTACAGGCGGTGAGGCGGTAAAGATTTCTTATATATAGTATCGAAATGG 257
QY 241 CACGAGAGTGTGAGAGGCGGCTTTCAGGGGCGCTGCGAGTGGATGCGAGGAGGAGCTG 300
DB 258 CACGAGAGTGTGAGAGGCGGCTTTCAGAGGCTGCTGCAAGTGGAAATGGAGAGCTG 317
QY 301 CAGGAGCTGTCCATCACTGTGCTCAACGCTCACTGTAACGACTCTGGCCTCTACACCTGC 360
DB 318 CAGGAGCTATCATCACTGTGCTCAATGCTCACTGTAATGCTGTGCGCTCTACACATGT 377
QY 361 AATGTGTCCGGAGTTGAGTTGAGGCGCATCGGCCCTTTGTGAAGACGACGCGGCTG 420
DB 378 AATGTGTCCAGGAGTTGAGTTGAGGCGCATCACCGGCCCTTTGTGAAGACGACCAAGACT 437
QY 421 ATCCCTTAAGAGTCAACGAGGAGGCTGAGAGAGCTTACCTCTGTGGTCTCAGAATC 480
DB 438 ATACCCCTGCGAGTCACTGAAGAGGCGGAGAGACTTACCTCCGCTGCGGAAATC 497
QY 481 ATGATGTACATCTTCTGCTTCTTCTCACCCTGTGGCTGCTCATCGAGATGATATATTCG 540
DB 498 ATGATGTACATCTTCTGCTTCTTCTCACCCTGTGGCTGCTTATGAGATGATATATTCG 557
QY 541 TACAGAAAGGTCTCAAAAGCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
DB 558 TACAGAAAGGTCTCTAAGGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 617
QY 601 ATCCCATCTGAGAACAGGAGAACTCTGCGGTACCACTGAGTGGAGGAGGAGGAGGAGGAG 647
DB 618 ATCCCTTTCAGAGAACAGGAGAACTCTGCTGTACCTGTGGAGGAGGAGGAGGAGGAGGAG 664
```



```
RESULT 10
RNO243395      2220 bp  mRNA  linear  ROD 29-JAN-2003
LOCUS          Rattus norvegicus mRNA for voltage-gated sodium channel beta-3
DEFINITION     subunit.
ACCESSION      AJ243395
VERSION        AJ243395.2 GI:7242802
KEYWORDS       scn3b gene; voltage-gated sodium channel beta-3 subunit.
SOURCE         Rattus norvegicus (Norway rat)
ORGANISM       Rattus norvegicus
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
               Rattus.
REFERENCE      1
AUTHORS        Morgan, K., Stevens, E.B., Shaw, B., Cox, P., Dixon, A.K., Lee, K.,
               Pincock, R.D., Hughes, J., Richardson, P.J., Mizuguchi, K. and
               Jackson, A.P.
TITLE          beta 3: an additional auxiliary subunit of the voltage-sensitive
JOURNAL        Proc. Natl. Acad. Sci. U.S.A. 97 (5), 2308-2313 (2000)
MEDLINE        20160948
PUBMED         10688874
REFERENCE      2
AUTHORS        Morgan, K.
TITLE          Direct Submission
JOURNAL        Submitted (28-JUN-1999) Morgan K., Biochemistry, University of
               Cambridge, Tennis Court Road, Cambridge, CB2 1QW, UNITED KINGDOM
REMARK         Revised by [3] (bases 1 to 2220)
REFERENCE      3
AUTHORS        Morgan, K.
TITLE          Direct Submission
JOURNAL        Submitted (09-MAR-2000) Morgan K., Biochemistry, University of
               Cambridge, Tennis Court Road, Cambridge, CB2 1QW, UNITED KINGDOM
COMMENT         On Mar 14, 2000 this sequence version replaced gi:7161888.
FEATURES       source
               1..2220
               /organism="Rattus norvegicus"
               /mol_type="mRNA"
               /db_xref="taxon:10116"
               /tissue_type="brain"
               /tissue_lib="lambda ZAP"
               1..2220
               /gene="scn3b"
               <1..362
               /gene="scn3b"
               363..1010
               /gene="scn3b"
               /function="modulates channel opening kinetics"
               /codon_start=1
               /product="voltage-gated sodium channel beta-3 subunit"
               /protein_id="CA876838.1"
               /db_xref="GI:7161889"
               /db_xref="GOA:Q9JUK0"
               /db_xref="UniProt/TREMBL:Q9JUK0"
               /translations="MPAFNRLPLSLVLIYVVRVCFVPCVVEPSTEAVQGNPKMLR
               CISCMEKEVATTVVWFVPRPGGKDFLIYVRNGHGVESPFQGLRWNGSKLOLD
               VSIIVNLVNDLSGLTCNVSRPFEPFHRPFVKTRLPLRVTEAGDFTSVSEI
               MNTILLVFLTLWLFIEIMYCYRKVSKAEAAQENASDYLAIPSENKENSVPVVEE"
               1011..>2220
               /gene="scn3b"
3'UTR
ORIGIN
Query Match      82.8%; Score 536.6; DB 10; Length 2220;
Best Local Similarity 89.3%; Pred. No. 4.1e-116;
Matches 578; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
QY 1 ATGCTGCTCCCTCAATAGATTGTTCCCTCGGCTTCTCTCGTGTCTATCTACTGGGTCAGT 60
Db 363 ATGCTGCTCCCTCAACAGATTGTTCCCTCGGCTTCTCTAGTGTCTATCTACTGGGTCA 422
QY 61 GTCTGCTTCCCTGTGTGTGGAAAGTGCCTCGGAGACGGAGCGCGTGCAGGGCAACCCC 120
```

```
423 GTCTGCTTCCCTGTGTGTGGAAAGTGCCTCGGAGACAGAGCGGTGCAGGGCAATCCC 482
121 ATGAAGCTGGCGTGCATCTCTCTCATGAGAGAGAGAGGTGGAGGCACACAGGTGGTG 180
483 ATGAAGCTGAGGTGCATCTCTCTCATGAGAGAGAGAGGTGGAGGCACACACTGTGGTG 542
181 GAATGGTTCTACAGGCCCGAGGCGCGTAAAGATTTCCTTATTACAGATATCGGAATGGC 240
543 GAGTGGTTCTACAGGCCCTGAGGCGCGTAAAGATTTCCTTATATATGAGTATCGGAATGGC 602
241 CACCAGAGGTGGAGAGCCCTTTTCAGGCGCGCTCGAGTGGAAATGSCAGCAAGGACCTG 300
603 CACCAGAGGTGGAGAGCCCTTTCCAAGGCGCTCGCAGTGGAAATGGGAGCAAGACCTG 662
301 CAGGAGGTGTCATCACTGCTGTCACAGTCACCTCTGAACGACTCTGGCCTCTACACTGC 360
663 CAGGAGGTATCCATCACTGCTGCTCAATGTCACCTTTGAAAGTACTCTGGCCTCTACATGC 722
361 AATGTGTCCCGGAGTTTGAAGTTTGAGGCGCATCGGCCCTTTTGTGAAGACGAGCGGCTG 420
723 AATGTGTCCAGGAGTTCGAAATTCGAGGCACACAGGCTTTTGTGAAGACGAGACTG 782
421 ATCCCTTAAGACTCACCGAGGAGGTGGAGAGACTTCACTCTGTGTTCTCAGAAATC 480
783 ATACCTTTGCGAGTCACTGAAGAGCGGAGAGAGACTTCACTCCGTGGTCTCGGAAATC 842
481 ATGATGTACATCTTCTGCTTCTCCCTCACCCTGTGGCTGCTCATCGAGATATATTCG 540
843 ATGATGTACATCTTCTGCTTCTCCCTCACCCTGTGGCTGTTTATTGAGATGATCTATTGC 902
541 TACAGAAAGGTCTCAAAAGCCGAGAGGAGCGCCCAAGAAACCGCTCTGACTACCTTGCC 600
903 TACAGAAAGGTCTCTAAGCCGAGAGGAGGAGCAGAGAAATCGCTGACTACCTTGCT 962
601 ATCCCATCTGAGAACCAAGAGAGAACTCTGCGGTACCAAGTGGAGGAATA 647
963 ATCCCTTCAGAGAACCAAGGAGAACTCTGTGTTACCTGTGGAGGAATA 1009

RESULT 11
AX048005      2632 bp  DNA  linear  PAT 15-DEC-2000
LOCUS          Sequence 22 from Patent WO0069912.
DEFINITION     AX048005
ACCESSION      AX048005
VERSION        AX048005.1 GI:11876883
KEYWORDS       Rattus sp.
SOURCE         Rattus sp.
ORGANISM       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
               Rattus.
REFERENCE      1
AUTHORS        Curtis, R.A.
TITLE          Gene encoding a sodium channel beta-3 subunit protein
JOURNAL        Patent: WO 0069912-A 22 23-NOV-2000;
               Millennium Pharmaceuticals, Inc. (US)
FEATURES       source
               1..2632
               /organism="Rattus sp."
               /mol_type="unassigned DNA"
               /db_xref="taxon:10118"
ORIGIN
Query Match      82.8%; Score 536.6; DB 6; Length 2632;
Best Local Similarity 89.3%; Pred. No. 4.1e-116;
Matches 578; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
QY 1 ATGCTGCTCCCTCAATAGATTGTTCCCTCGGCTTCTCTCGTGTCTATCTACTGGGTCA 60
Db 78 ATGCTGCTCCCTCAACAGATTGTTCCCTCGGCTTCTCTAGTGTCTATCTACTGGGTCA 137
QY 61 GTCTGCTTCCCTGTGTGTGGAAAGTGCCTCGGAGACGGAGCGCGTGCAGGGCAACCCC 120
```


Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. Rattus.

1 Curtis, R.A.
Gene encoding a sodium channel beta-3 subunit protein
Patent: WO 0069912-A 1 23-NOV-2000;
Millennium Pharmaceuticals, Inc. (US)
Location/Qualifiers
1..3108
/organism="Rattus sp."
/mol_type="unassigned DNA"
/db_xref="taxon:10118"

ORIGIN
Query Match 82.8%; Score 536.6; DB 6; Length 3108;
Best Local Similarity 89.3%; Pred. No. 4.1e-116;
Matches 578; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 1 ATGCCTGCTTCAATAGATTGTTCCCTGCTTCTCTGCTTATCTACTGGTCACT 60
Db 78 ATGCCTGCTTCAACAGATTGTTCCCTGCTTCTCTGCTTATCTACTGGTCACT 137
QY 61 GTCTGCTTCCCTGCTGTTGGAAGTCCCTCGAGACGAGCGCGTGCAGGCAACCC 120
Db 138 GTCTGCTTCCCTGCTGTTGGAAGTCCCTCGAGACGAGCGCGTGCAGGCAACCC 197
QY 121 ATGAAGTGGCTGCATCTCTCCATGAAGAGAGAGAGTGGAGGCCACACGGTGGT 180
Db 198 ATGAAGTGGAGTGCATCTCTCCATGAAGAGAGAGAGTGGAGGCCACACGGTGGT 257
QY 181 GAATGTTTACAGGCGGCGGTAAGATTTCCTTATTTACAGATATCGGAATGGC 240
Db 258 GAGTGGTCTACAGGCTGAGGCGGTAAGATTTCCTTATATATGATATCGGAATGGC 317
QY 241 CACAGAGTGGAGAGCCCTTTACAGGCGGCTGCGAGTGAATGCGAAGAGACCTG 300
Db 318 CACAGAGTGGAGAGCCCTTTACAGGCGGCTGCGAGTGAATGCGAAGAGACCTG 377
QY 301 CAGAGTGTCCATCACTGCTCAAGTCACTGCACTGCACTGCACTGCACTGCACTG 360
Db 378 CAGAGTGTCCATCACTGCTCAAGTCACTGCACTGCACTGCACTGCACTGCACTG 437
QY 361 AATGTGTCCGCGAGTTGAGTTGAGCGGCGATCGGCGCTTTGTGAAGACGAGCGGCTG 420
Db 438 AATGTGTCCGCGAGTTGAGTTGAGCGGCGATCGGCGCTTTGTGAAGACGAGCTG 497
QY 421 ATCCCTTAAAGTCAACGAGGCGTGGAGAGCTTCACTGCACTGCTGCTGCTGCTGCTG 480
Db 498 ATACCTTTGCGAGTCACTGAAGAGGCGGAGAGACTTCACTTCCGTTGCTCGGAATC 557
QY 481 ATGATGTACATCTTCTGCTTCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Db 558 ATGATGTACATCTTCTGCTTCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 617
QY 541 TACAGAAAGTCTCAAAAGCCGAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
Db 618 TACAGAAAGTCTCAAAAGCCGAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 677
QY 601 ATCCCATCTGAGAACAGGAGAACTCTGCGGTACCAAGTGGAGGAATA 647
Db 678 ATCCCTTACAGAGAACAGGAGAACTCTGCGGTACCAAGTGGAGGAATA 724

RESULT 14
AKI73115
LOCUS Mus musculus mRNA for mKIAA1158 protein. ROD 28-JUL-2004
DEFINITION AKI73115
ACCESSION AKI73115
VERSION AKI73115.1 GI:50510814
KEYWORDS FLI_CDNA.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE AUTHORS

TITLE

JOURNAL REFERENCE AUTHORS

TITLE JOURNAL

COMMENT

FEATURES source

gene

CDS

ORIGIN

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. Rattus.

1 Okazaki, N., Kikuno, R.F., Ohara, R., Inamoto, S., Koseki, H., Hiraoka, S., Saga, Y., Seino, S., Nishimura, M., Kaisho, T., Hoshino, K., Kitamura, H., Nagase, T., Ohara, O. and Koga, H.
Prediction of the Coding Nucleotide Sequences of Mouse Homologues of KIAA
Gene: IV. The Complete Nucleotide Sequences of 500 Mouse
KIAA-Homologous cDNAs Identified by Screening of Terminal Sequences
of cDNA Clones Randomly Sampled from Size-Fractionated Libraries
DNA Res. 11, 205-218 (2004)
2 (bases 1 to 4025)
Okazaki, N., Kikuno, R.F., Nagase, T., Ohara, O. and Koga, H.
Direct Submission
Submitted (19-MAY-2004) Hisashi Koga, Kazusa DNA Research
Institute, Laboratory for Genome Informatics; 2-6-7
Kazusa-kamatari, Kisarazu, Chiba 292-0818, Japan
(E-mail:mouse@kazusa.or.jp, Tel:81-438-52-3919, Fax:81-438-52-3918)
The CREATE program supported by Japan science and technology
corporation; cDNA full insert sequencing; Kazusa DNA Research
Institute; cDNA library construction, clone selection and 5'- &
3'-end one pass sequencing.
Location/Qualifiers
1..4025
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="mfj07349"
/notes="vector:modified pBC SK+"
1..4025
/genes="mKIAA1158"
/genes="mKIAA1158"
/note="CDS is predicted by in silico analysis. Start codon
is not identified."
/evidence=not_experimental
/product="mKIAA1158 protein"
/protein_id="BAD32393.1"
/db_xref="GI:50510815"
/translation="SSTHSVAQTQTAGVTRPAGEACPAVEARDSESRWIEVAVPRGSE
SRNHLREKVPAPFNRLPLASLVLYVVRVCFVPEVSEFSEAVQNSMKRCSIM
KREEVATVWFVPEGGKDFLIYFNRHGVESPFQRLQWMSKDLQVSTIV
LNVTLNDSGLYTVNSRFEFPAHPRFKTRILPLRVTEAGEDFTSVVSEIMMYIL
LVFLTLWLFMIYCYRKVSKAEAAQENASDYLAIPTSENKENVVPEVE"

Query Match 82.8%; Score 536.6; DB 10; Length 4025;
Best Local Similarity 89.3%; Pred. No. 4.1e-116;
Matches 578; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 1 ATGCCTGCTTCAATAGATTGTTCCCTGCTTCTCTGCTTATCTACTGGTCACT 60
Db 290 ATGCCTGCTTCAACAGATTGTTCCCTGCTTCTCTGCTTATCTACTGGTCACT 349
QY 61 GTCTGCTTCCCTGCTGTTGGAAGTCCCTCGAGACGAGCGCGTGCAGGCAACCC 120
Db 350 GTCTGCTTCCCTGCTGTTGGAAGTCCCTCGAGACGAGCGCGTGCAGGCAATTC 409
QY 121 ATGAAGTGGCTGCATCTCTGCTGATGAAGAGAGAGTGGAGGCCACACGGTGGT 180
Db 410 ATGAAGTGGAGTGCATCTCTGCTGATGAAGAGAGTGGAGGCCACACGGTGGT 469
QY 181 GAATGTTTACAGGCGGCGGTAAGATTTCCTTATTTACAGTATCGGAATGGC 240
Db 470 GAGTGGTCTTACAGGCGGTAAGATTTCCTTATATATGATATCGGAATGGC 529
QY 241 CACAGAGTGGAGAGCCCTTTACAGGCGGCTGCGAGTGAATGCGAAGAGACCTG 300
Db 530 CACAGAGTGGAGAGCCCTTTACAGGCTGCTGCGAGTGAATGCGAAGAGACCTG 589
QY 301 CAGGAGTGTCCATCACTGCTGCTCAACGCTCACTGCTGCTGCTGCTGCTGCTGCTGCT 360

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2005, 06:13:29 ; Search time 470.712 Seconds
(without alignments)
8149.338 Million cell updates/sec

Title: US-09-977-579-4_COPY_376_1023

Perfect score: 648

Sequence: 1 atgcctgccttcattagatt.....cgggtaccagtgagggaatag 648

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|---------------------|
| 1 | 648 | 100.0 | 1261 | 3 | AAC67837 Human bet |
| 2 | 648 | 100.0 | 1261 | 10 | ACF57870 Human SCN |
| 3 | 646.4 | 99.8 | 1261 | 10 | ADB78651 Human ion |
| 4 | 646.4 | 99.8 | 1510 | 4 | AAF84146 Human nov |
| 5 | 646.4 | 99.8 | 4052 | 6 | ABA93727 Human sig |
| 6 | 585 | 90.3 | 1045 | 4 | AAK52345 Human sig |
| 7 | 570.2 | 88.0 | 978 | 4 | AAN98320 Human EST |
| 8 | 570.2 | 88.0 | 978 | 13 | ADSI11487 Human the |
| 9 | 536.6 | 82.8 | 2220 | 3 | AAC67836 Rat beta3 |
| 10 | 536.6 | 82.8 | 2632 | 5 | AAC90602 Rat sodiu |
| 11 | 536.6 | 82.8 | 3108 | 5 | AAC90600 Rat sodiu |
| 12 | 534.8 | 82.5 | 1195 | 5 | AAS86764 DNA encod |
| 13 | 534.6 | 82.5 | 645 | 5 | AAC90601 Rat sodiu |
| 14 | 504 | 77.8 | 953 | 13 | ADSI10151 Human the |
| 15 | 422 | 65.1 | 471 | 2 | AAV86895 EST clone |
| 16 | 392.6 | 60.6 | 3531 | 5 | AAS86763 DNA encod |
| 17 | 224.4 | 34.6 | 621 | 5 | AAS86762 DNA encod |
| 18 | 148.2 | 22.9 | 657 | 5 | AAC90603 Rabbit so |
| 19 | 143.8 | 22.2 | 1335 | 10 | ACF57868 Human SCN |
| 20 | 143.8 | 22.2 | 1335 | 12 | ADQ19428 Human sof |

| | | | | | |
|------|-------|------|------|----|---------------------|
| 21 | 142.2 | 21.9 | 1414 | 8 | ABZ23837 Human vol |
| 22 | 142.2 | 21.9 | 1414 | 10 | ADB78640 Human ion |
| 23 | 142.2 | 21.9 | 1414 | 10 | ADB78641 Human ion |
| c 24 | 141 | 21.8 | 407 | 4 | ABA08942 Human vol |
| c 25 | 141 | 21.8 | 407 | 4 | AAK53329 Human pol |
| 26 | 138.6 | 21.4 | 1490 | 6 | ABK63747 Rat sequ |
| 27 | 138.6 | 21.4 | 1490 | 10 | ADB52855 Primary r |
| 28 | 138.6 | 21.4 | 1490 | 12 | ADO09498 Rat sodiu |
| 29 | 81.4 | 12.6 | 974 | 6 | AAD29622 Human bet |
| c 30 | 73.8 | 11.4 | 358 | 10 | ACD98234 Human col |
| 31 | 73 | 11.3 | 850 | 4 | AAF57675 Rat sodiu |
| 32 | 73 | 11.3 | 850 | 4 | AAF30669 Sodium ch |
| 33 | 72.6 | 11.2 | 855 | 10 | ADGI15041 Human SEC |
| 34 | 49 | 7.6 | 243 | 12 | ACH83019 Human gen |
| 35 | 49 | 7.6 | 569 | 12 | ACH69319 Human gen |
| 36 | 48.6 | 7.5 | 182 | 12 | ADM66944 Human adl |
| 37 | 48.6 | 7.5 | 3583 | 12 | ADQ23651 Human sof |
| 38 | 48 | 7.4 | 842 | 12 | ADM66943 Murine ad |
| 39 | 47 | 7.3 | 509 | 4 | AAI93385 Human pol |
| 40 | 45.4 | 7.0 | 1645 | 3 | AAA88851 Maize ger |
| 41 | 45.4 | 7.0 | 1645 | 12 | ADI36508 Corn gera |
| 42 | 44.8 | 6.9 | 921 | 13 | ADT44612 Bacterial |
| c 43 | 44 | 6.8 | 53 | 12 | ADQ07690 PCR prime |
| 44 | 44 | 6.8 | 60 | 12 | ADQ07691 PCR prime |
| 45 | 43.8 | 6.8 | 1434 | 12 | ADI45691 Corn isop |

ALIGNMENTS

RESULT 1

AAC67837

ID AAC67837 standard; cDNA; 1261 BP.

XX AAC67837;

XX 15-FEB-2001 (first entry)

XX Human beta3 cDNA.

XX Human; beta sub-unit; beta3; analgesic; anticonvulsant;

KW cerebrotective; vasotropic; cardiac; nootropic; cytostatic;

KW dermatological; gene therapy; voltage-gated sodium channel; pain;

KW epilepsy; stroke; ischaemia; heart disease; Jacobsen Syndrome;

KW familial nonchromaffin paraganglioma; phenylketonuria;

KW Charcot Marie Tooth disease; ss.

XX Homo sapiens.

XX WO200063367-A1.

XX 26-OCT-2000.

XX 24-FEB-2000; 2000WO-EP001783.

XX 15-APR-1999; 99US-0129473P.

XX (WARN) WARNER LAMBERT CO.

XX (UYCA-) UNIV CAMBRIDGE TECH SERVICES LTD.

XX Cox P, Dixon A, Jackson A, Morgan K;

XX WPI; 2000-665241/64..

XX P-PSDB; AAB36002.

XX Novel nucleic acids encoding a beta-3 subunit from a voltage-gated sodium

PT channel, and their corresponding polypeptides, useful for detecting and

PT treating sodium channel-associated conditions, e.g. pain, epilepsy and

PT stroke.

XX Claim 10; Page 70-71; 88pp; English.

XX The present sequence is given in the claims of a specification relating

CC

CC to a novel family of beta sub-unit proteins from a voltage-gated sodium
CC channel. Human and rat beta sub-units, which have been collectively
CC identified as beta3, have been isolated. The polynucleotides and
CC polypeptides are useful for screening for agonists and antagonists of
CC sodium channels. The agonists, antagonists, proteins and nucleic acids
CC may be used diagnosing of treating diseases or conditions associated with
CC voltage-gated sodium channels, e.g. pain, epilepsy, stroke, ischaemia,
CC heart disease, Jacobsen Syndrome, Familial Nonchromaffin Paraganglioma,
CC Phenylketonuria and Charcot Marie Tooth disease
XX
XX
SQ Sequence 1261 BP; 263 A; 366 C; 372 G; 260 T; 0 U; 0 Other;

Query Match 100.0%; Score 648; DB 3; Length 1261;
Best Local Similarity 100.0%; Pred. No. 3.9e-165;
Matches 648; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCCTGCTTCAATAGATTGTTTCCCTGGCTTCTCTGCTTATCTACTGGGTCACT 60
Db |||||
QY 376 ATGCCTGCTTCAATAGATTGTTTCCCTGGCTTCTCTGCTTATCTACTGGGTCACT 435
Db |||||
QY 61 GTCCTGCTCCCTGTGTGTGGAGTGCCTCGAGACGGAGCCGTGAGGCAACCC 120
Db |||||
QY 436 GTCCTGCTCCCTGTGTGTGGAGTGCCTCGAGACGGAGCCGTGAGGCAACCC 495
Db |||||
QY 121 ATGAAGCTGGCTGCATCTCTCTCATGAGAGAGAGGTGAGGCCACCGTGGTG 180
Db |||||
QY 496 ATGAAGCTGGCTGCATCTCTCTCATGAGAGAGAGGTGAGGCCACCGTGGTG 555
Db |||||
QY 181 GAATGCTTCTACAGCCGAGGCGGTAAAGATTCTCTTATTACAGTATCGAATGCG 240
Db |||||
QY 556 GAATGCTTCTACAGCCGAGGCGGTAAAGATTCTCTTATTACAGTATCGAATGCG 615
Db |||||
QY 241 CACAGAGGTGAGAGCCCTTTCAGGGCGCTCGAGTGAATGGCAGCAAGCACTG 300
Db |||||
QY 616 CACAGAGGTGAGAGCCCTTTCAGGGCGCTCGAGTGAATGGCAGCAAGCACTG 675
Db |||||
QY 301 CAGGAGGTGCATCTGCTCAAGTCACTCAAGTCACTCAAGTCACTCAAGTCACT 360
Db |||||
QY 676 CAGGAGGTGCATCTGCTCAAGTCACTCAAGTCACTCAAGTCACTCAAGTCACT 735
Db |||||
QY 361 AATGCTGCTCCGGAGTTTGAAGTGAAGGCGCATCGGCCCTTTGTGAAGACGACGCGCTG 420
Db |||||
QY 736 AATGCTGCTCCGGAGTTTGAAGTGAAGGCGCATCGGCCCTTTGTGAAGACGACGCGCTG 795
Db |||||
QY 421 ATCCCTTAAAGTCAACGAGAGGTGAGAGGATTCACCTCTGTGTCTCAGAAATC 480
Db |||||
QY 796 ATCCCTTAAAGTCAACGAGAGGTGAGAGGATTCACCTCTGTGTCTCAGAAATC 855
Db |||||
QY 481 ATGATGTACATCTTCTGTGCTTCTCCTCACCCTGTGGCTGCTCATCGAGATGATATTCG 540
Db |||||
QY 856 ATGATGTACATCTTCTGTGCTTCTCCTCACCCTGTGGCTGCTCATCGAGATGATATTCG 915
Db |||||
QY 541 TACAGAAAGTCTCAAAAGCCGAGAGGAGCCCAAGAAACCGCTCTGACTACCTTTGCC 600
Db |||||
QY 916 TACAGAAAGTCTCAAAAGCCGAGAGGAGCCCAAGAAACCGCTCTGACTACCTTTGCC 975
Db |||||
QY 601 ATCCCATCTGAGAACAGAGAACTCTCTGGGTACCAAGTGGAGGAATAG 648
Db |||||
QY 976 ATCCCATCTGAGAACAGAGAACTCTCTGGGTACCAAGTGGAGGAATAG 1023
Db |||||

RESULT 2
ACF57870
ID ACF57870 standard; cDNA; 1261 BP.
XX
AC ACF57870;
XX
DT 15-JAN-2004 (first entry)
XX
DE Human SCN3B protein encoding cDNA.
XX
XX SCN1A; sodium channel type 1 alpha-subunit; anticonvulsant; analgesic;
KW neuroprotective; anesthetic; cyostatic; cerebroprotective; cardiant;
KW

hypotensive; gene therapy; SCN3B; human; gene; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
FT CDS 376..1023
FT /tag= a
FT /product= "SCN3B"
XX
XX WO2003072751-A2.
XX
XX 04-SEP-2003.
XX 25-FEB-2003; 2003WO-US006010.
XX 25-FEB-2002; 2002US-0359382P.
XX (UYVA-) UNIV VANDERBILT.
XX George AL, Lessin C;
XX WPI; 2003-712725/67.
XX P-PSDB; ABR83183.
XX Recombinantly expressed sodium channel type 1 alpha subunit, useful in
XX screening for modulators, for treating e.g. epilepsy.
XX Disclosure; Page 145-147; 176pp; English.
XX The invention relates to a recombinantly expressed and isolated human
XX SCN1A (sodium channel type 1 alpha-subunit) (I). (I), optionally
XX incorporated into a cell, is used to screen for specific modulators,
XX potentially useful as anticonvulsant, antiepileptic, neuroprotective,
XX analgesic and/or anesthetic agents, e.g. for treating severe myoclonic
XX epilepsy of infancy, stroke, cardiac arrest, hyperkalemic paralysis,
XX motor endplate diseases, hypertension, congestive heart failure and
XX muscular dystrophy also to treat cancer (SCN1A is expressed in prostatic
XX and metastatic cancer cell lines). These activities can also be provided
XX by gene therapy vectors that express (I) or the modulators. The
XX modulators, also antibodies directed against (I), are used to detect
XX sodium channel polypeptides. The present sequence represents a human
XX SCN3B protein encoding cDNA
XX
SQ Sequence 1261 BP; 263 A; 366 C; 372 G; 260 T; 0 U; 0 Other;

Query Match 100.0%; Score 648; DB 10; Length 1261;
Best Local Similarity 100.0%; Pred. No. 3.9e-165;
Matches 648; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCCTGCTTCAATAGATTGTTTCCCTGGCTTCTCTGCTTATCTACTGGGTCACT 60
Db |||||
QY 376 ATGCCTGCTTCAATAGATTGTTTCCCTGGCTTCTCTGCTTATCTACTGGGTCACT 435
Db |||||
QY 61 GTCCTGCTCCCTGTGTGTGGAGTGCCTCGAGACGGAGCCGTGAGGCAACCC 120
Db |||||
QY 436 GTCCTGCTCCCTGTGTGTGGAGTGCCTCGAGACGGAGCCGTGAGGCAACCC 495
Db |||||
QY 121 ATGAAGCTGGCTGCATCTCTCTCATGAGAGAGAGGTGAGGCCACCGTGGTG 180
Db |||||
QY 496 ATGAAGCTGGCTGCATCTCTCTCATGAGAGAGAGGTGAGGCCACCGTGGTG 555
Db |||||
QY 181 GAATGCTTCTACAGCCGAGGCGGTAAAGATTCTCTTATTACAGTATCGAATGCG 240
Db |||||
QY 556 GAATGCTTCTACAGCCGAGGCGGTAAAGATTCTCTTATTACAGTATCGAATGCG 615
Db |||||
QY 241 CACAGAGGTGAGAGCCCTTTTCAGGGCGCTCGAGTGAATGGCAGCAAGCACTG 300
Db |||||
QY 616 CACAGAGGTGAGAGCCCTTTTCAGGGCGCTCGAGTGAATGGCAGCAAGCACTG 675
Db |||||
QY 301 CAGGAGGTGCATCTGCTCAAGTCACTCTGAACGACTCTGGCTCTACACCTGC 360
Db |||||
QY 676 CAGGAGGTGCATCTGCTCAAGTCACTCTGAACGACTCTGGCTCTACACCTGC 735
Db |||||

QY 361 AATGTGTCCCGGAGTTTGAGTTTGAGCGCGCATCGGCCCTTTGTGAAGACGACGCGCTG 420
 Db |||||
 736 AATGTGTCCCGGAGTTTGAGTTTGAGCGCGCATCGGCCCTTTGTGAAGACGACGCGCTG 795
 QY 421 ATCCCTTAAGTACACGAGAGGCTGGAGAGACTTCACTCTGTGTGCTCAGAAATC 480
 Db |||||
 796 ATCCCTTAAGTACACGAGAGGCTGGAGAGACTTCACTCTGTGTGCTCAGAAATC 855
 QY 481 ATGATGTACATCTTCTGTGCTTCTCCTCACCTGTGGCTGCTCATCGAGATGATATATTC 540
 Db |||||
 856 ATGATGTACATCTTCTGTGCTTCTCCTCACCTGTGGCTGCTCATCGAGATGATATATTC 915
 QY 541 TACAGAAAGGTCTCAAAAGCCGAGAGCGAGCGCCCAAGAAACCGCTCTGACTACCTTGCC 600
 Db |||||
 916 TACAGAAAGGTCTCAAAAGCCGAGAGCGAGCGCCCAAGAAACCGCTCTGACTACCTTGCC 975
 QY 601 ATCCCATCTGAGACAGGAGAACTCTGCGGTACCACTGTGAGGAATAG 648
 Db |||||
 976 ATCCCATCTGAGACAGGAGAACTCTGCGGTACCACTGTGAGGAATAG 1023

RESULT 3

ADB78651

ID ADB78651 standard; cDNA; 1261 BP.

XX

AC ADB78651;

XX

DT 04-DEC-2003 (first entry)

XX

DE Human ion channel subunit cDNA mutant SCN1Ar exon 1 SEQ ID NO:22.

XX

SW ss; gene; mutant; ion channel; ion channel subunit; ICS; nootropic;
 KW neuroprotective; inotropic; antipyretic; antiarrhythmic; antimigraine;
 KW antidepressant; antiparkinsonian; neuroleptic; tranquiliser; analgesic;
 KW nephrotropic; antidiabetic; ophthalmological; epilepsy;
 KW ion channel dysfunction; human.

XX

OS Synthetic.

OS

OS Homo sapiens.

XX

PN WO2003008574-A1.

XX

PD 30-JAN-2003.

XX

XX 08-JUL-2002; 2002WO-AU000910.

XX

XX 18-JUL-2001; 2001AU-00006452.

PR

PR 05-MAR-2002; 2002AU-00000910.

PR

PR 13-MAY-2002; 2002AU-00002292.

XX

XX (BION-) BIONOMICS LTD.

PA

PA (WALL-) WALLACE R W.

PA

XX Mulley JC, Harkin LA, Dibbens LM, Phillips HA, Heron SE;

PI

PI Berkovic SF, Scheffer IE;

PI

DR WPI; 2003-239332/23.

XX

XX Identifying predisposition to an ion channel dysfunction, such as

PT

PT periodic paralysis, cardiac arrhythmias, migraine, Alzheimer's disease,
 PT schizophrenia, anxiety and depression, by detecting encoding-gene
 PT mutation events.

XX

PS Claim 6; SEQ ID NO 22; 106pp; English.

XX

CC The invention relates to a novel method for identifying a subject
 CC predisposed to a disorder associated with ion channel dysfunction. The
 CC method comprises ascertaining if at least one of the genes encoding ion
 CC channel subunits (ICS) has undergone a mutation event so that a cDNA
 CC derived from the subject has any of 134 nucleotide sequences. The method
 CC of the invention has nootropic, neuroprotective, inotropic, antipyretic,
 CC antiarrhythmic, antimigraine, antidepressant, antiparkinsonian,
 CC neuroleptic, tranquiliser, analgesic, nephrotropic, antidiabetic, and

CC ophthalmological activity. A polynucleotide of the invention acts as an
 CC ion channel agonist, or ion channel antagonist. The methods, isolated
 CC nucleic acids, polypeptides, antibody, genetically agonist, antagonist or
 CC modulator of an ion channel, cells and genetically modified non-human
 CC animal, are useful for the diagnosis and treatment of epilepsy and/or a
 CC disorder associated with ion channel dysfunction, such as hyper- or hypo-
 CC kalemic periodic paralysis, myotonias, malignant hyperthermia,
 CC myasthenia, cardiac arrhythmias, episodic ataxia, migraine, Alzheimer's
 CC disease, Parkinson's disease, schizophrenia, hyperkplexia, anxiety,
 CC depression, phobic obsessive symptoms, neuropathic pain, inflammatory
 CC pain, chronic/acute pain, Bartter's syndrome, polycystic kidney disease,
 CC Dent's disease, hyperinsulinaemic hypoglycaemia of infancy, cystic
 CC fibrosis, congenital stationary night blindness and total colour
 CC blindness. The present sequence represents a mutant cDNA of the
 CC invention. The sequence data for this patent is not represented in the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pat_sequences.

XX

SQ Sequence 1261 BP; 263 A; 365 C; 372 G; 261 T; 0 U; 0 Other;

Query Match

Best Local Similarity 99.8%; Score 646.4; DB 10; Length 1261;

Matches 647; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGCTGCTTCAATAGATTGTTTCCCTGGCTTCTCTCGTGTATCTACTGGGTGAGT 60

Db

376 ATGCTGCTTCAATAGATTGTTTCCCTGGCTTCTCTCGTGTATCTACTGGGTGAGT 435

QY

61 GTCTGCTTCTCGTGTGTGGAAATGCTCTCGAGACGAGGCGGTGAGGGCAACCCC 120

Db

436 GTCTGCTTCTCGTGTGTGGAAATGCTCTCGAGACGAGGCGGTGAGGGCAACCCC 495

QY

121 ATCAAGCTGCGTGCATCTCTCATGAGAGAGAGGTGAGGCGCCACCGTGTG 180

Db

496 ATGAAGCTGCGTGCATCTCTCATGAGAGAGAGGTGAGGCGCCACCGTGTG 555

QY

181 GAATGTTTCTACAGGCGCGGTAAAGATTTCCTTATTTACGAGTATCGGAATGCG 240

Db

556 GAATGTTTCTACAGGCGCGGTAAAGATTTCCTTATTTACGAGTATCGGAATGCG 615

QY

241 CACGAGAGGTGGAGAGCGCTTTTCAAGGCGCGCTGTCAGTGGAAATGCGCAGGACCTG 300

Db

616 CACGAGAGGTGGAGAGCGCTTTTCAAGGCGCGCTGTCAGTGGAAATGCGCAGGACCTG 675

QY

301 CAGGAGGTGTCATCACTGTCTCAACGTCACCTCTGAACGACTCTGSCCTCTACACCTGC 360

Db

676 CAGGAGGTGTCATCACTGTCTCAACGTCACCTCTGAACGACTCTGSCCTCTACACCTGC 735

QY

361 AATGTGTCCCGGAGTTTGAAGTTCGAGCGCATCGGCCCTTTTGTGAAGACGACGCGCTG 420

Db

736 AATGTGTCCCGGAGTTTGAAGTTCGAGCGCATCGGCCCTTTTGTGAAGACGACGCGCTG 795

QY

421 ATCCCTTAAGAGTCAACGAGAGGCTGGAGAGACTTCACTCTGTGTGCTCAGAAATC 480

Db

796 ATCCCTTAAGAGTCAACGAGAGGCTGGAGAGACTTCACTCTGTGTGCTCAGAAATC 855

QY

481 ATGATGTACATCTTCTGTGCTTCTCCTCACCTGTGCTGCTCATCGAGATGATATATTC 540

Db

856 ATGATGTACATCTTCTGTGCTTCTCCTCACCTGTGCTGCTCATCGAGATGATATATTC 915

QY

541 TACAGAAAGGTCTCAAAAGCCGAGAGCGCCCAAGAAACCGCTCTGACTACCTTGCC 600

Db

916 TACAGAAAGGTCTCAAAAGCCGAGAGCGCCCAAGAAACCGCTCTGACTACCTTGCC 975

QY

601 ATCCCATCTGAGACAGGAGAACTCTGCGGTACCACTGTGAGGAATAG 648

Db

976 ATCCCATCTGAGACAGGAGAACTCTGCGGTACCACTGTGAGGAATAG 1023

XX

RESULT 4

AAF84146

ID AAF84146 standard; cDNA; 1510 BP.

XX

Matches 647; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGCTGCTCCCTCAATAGATGTTTCCCTGCTTCTCTCGTGCCTTACTACTGGTCACT 60
 DB 804 ATGCTGCTCCCTCAATAGATGTTTCCCTGCTTCTCTCGTGCCTTACTACTGGTCACT 863
 QY 61 GTCTGCTTCCCTGTGTGTGTGAAGTGCCTTCGAGAGCGAGCGCTGCAGGGCAACCCC 120
 DB 864 GTCTGCTTCCCTGTGTGTGTGAAGTGCCTTCGAGAGCGAGCGCTGCAGGGCAACCCC 923
 QY 121 ATGAAGCTGCTGCTCACTCTCTGATGAAGAGAGAGAGGTGAGGCCACACCGTGGTG 180
 DB 924 ATGAAGCTGCTGCTCACTCTCTGATGAAGAGAGAGAGGTGAGGCCACACCGTGGTG 983
 QY 181 GAATGGTTCACAGCCCGAGGGCGGTAAAGATTTCTTTATACAGATATCGGAATGGC 240
 DB 984 GAATGGTTCACAGCCCGAGGGCGGTAAAGATTTCTTTATACAGATATCGGAATGGC 1043
 QY 241 CACGAGAGGTGGAGAGCCCTTTTCAGGGGCGCTCGAGTGGATGGCAGCAAGACCTG 300
 DB 1044 CACGAGAGGTGGAGAGCCCTTTTCAGGGGCGCTCGAGTGGATGGCAGCAAGACCTG 1103
 QY 301 CAGAGCTGCTCACTACTGCTCAAGTCACTCTGAAGAGCTTGGCCTTACACCTGC 360
 DB 1104 CAGAGCTGCTCACTACTGCTCAAGTCACTCTGAAGAGCTTGGCCTTACACCTGC 1163
 QY 361 AATGTGTCCTGGAGTTGAGTTTGAGCGGCATCGGCCCTTTGGAAGAGCGCGGCTG 420
 DB 1164 AATGTGTCCTGGAGTTGAGTTTGAGCGGCATCGGCCCTTTGGAAGAGCGCGGCTG 1223
 QY 421 ATCCCTTAAAGAGTCAACGAGGAGGCTGGAGAGCTTACCTCTGTGGTCTCAGAATC 480
 DB 1224 ATCCCTTAAAGAGTCAACGAGGAGGCTGGAGAGCTTACCTCTGTGGTCTCAGAATC 1283
 QY 481 ATGATGTACATCTTCTGTGCTTCTCCTCACTGTGCTGCTCATCGAGATGATATATGC 540
 DB 1284 ATGATGTACATCTTCTGTGCTTCTCCTCACTGTGCTGCTCATCGAGATGATATATGC 1343
 QY 541 TACAGAAAGGTCTCAAAAGCGGAGAGCGAGCCCAAGAAACGGCTCTGACTACCTTGGC 600
 DB 1344 TACAGAAAGGTCTCAAAAGCGGAGAGCGAGCCCAAGAAACGGCTCTGACTACCTTGGC 1403
 QY 601 ATCCCATCTGAGAACAGGAGAACTCTGCGGTACCACTGGAGGAATAG 648
 DB 1404 ATCCCATCTGAGAACAGGAGAACTCTGCGGTACCACTGGAGGAATAG 1451

RESULT 6
 AAK52345
 ID AAK52345 standard; cDNA; 1045 BP.
 AC AAK52345;
 XX
 XX
 DE 06-NOV-2001 (first entry)
 XX
 XX Human polynucleotide SEQ ID NO 890.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation; ss.
 XX
 OS Homo sapiens.
 XX
 PN W0200157190-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001WO-US004098.
 XX
 PR 03-FEB-2000; 2000US-00496914.
 PR 27-APR-2000; 2000US-00560875.
 PR 20-JUN-2000; 2000US-00598075.

PR 19-JUL-2000; 2000US-00620325.
 PR 01-SEP-2000; 2000US-00654936.
 PR 15-SEP-2000; 2000US-00663561.
 PR 20-OCT-2000; 2000US-00693325.
 PR 30-NOV-2000; 2000US-00728422.
 XX (HYSE-) HYSEQ INC.
 XX
 XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y,
 PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 XX WPI; 2001-476283/51.
 DR P-PSDB; AAM79212.
 XX
 XX Nucleic acids encoding polypeptides with cytokine-like activities, useful
 PT in diagnosis and gene therapy.
 PT
 XX Claim 1; Page 2934-2935; 6221pp; English.
 XX
 XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
 CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
 CC sequence listing were missing at the time of publication
 XX
 SQ Sequence 1045 BP; 222 A; 289 C; 314 G; 220 T; 0 U; 0 Other;
 Query Match 90.3%; Score 585; DB 4; Length 1045;
 Best Local Similarity 99.2%; Pred. No. 4.2e-148;
 Matches 588; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 ATGCTGCTCCCTCAATAGATGTTTCCCTGCTTCTCTCGTGCCTTACTACTGGTCACT 60
 DB 302 ATGCTGCTCCCTCAATAGATGTTTCCCTGCTTCTCTCGTGCCTTACTACTGGTCACT 361
 QY 61 GTCTGCTTCCCTGTGTGTGTGAAGTGCCTTCGAGAGCGAGCGCTGCAGGGCAACCCC 120
 DB 362 GTCTGCTTCCCTGTGTGTGTGAAGTGCCTTCGAGAGCGAGCGCTGCAGGGCAACCCC 421
 QY 121 ATGAAGCTGCTGCTCATCTCTGATGAAGAGAGAGAGGTGGAGGCCACACCGTGGTG 180
 DB 422 ATGAAGCTGCTGCTCATCTCTGATGAAGAGAGAGAGGTGGAGGCCACACCGTGGTG 481
 QY 181 GAATGGTTCACAGCCCGAGGGCGGTAAAGATTTCTTTATACAGATATCGGAATGGC 240
 DB 482 GAATGGTTCACAGCCCGAGGGCGGTAAAGATTTCTTTATACAGATATCGGAATGGC 541
 QY 241 CACGAGAGGTGGAGAGCCCTTTTCAGGGGCGCTCGAGTGGATGGCAGCAAGACCTG 300
 DB 542 CACGAGAGGTGGAGAGCCCTTTTCAGGGGCGCTCGAGTGGATGGCAGCAAGACCTG 601
 QY 301 CAGGAGGTGCTCATCACTGTGCTCAACGTCACTCTGNAACGACTCTGCGCTTACACCTGC 360
 DB 602 CAGGAGGTGCTCATCACTGTGCTCAACGTCACTCTGNAACGACTCTGCGCTTACACCTGC 661
 QY 361 AATGTGTCCTGGAGTTTGAAGTGGCGGCATCGGCCCTTTGTGAAGAGCGCGGCTG 420
 DB 662 AATGTGTCCTGGAGTTTGAAGTGGCGGCATCGGCCCTTTGTGAAGAGCGCGGCTG 721
 QY 421 ATCCCTTAAAGAGTCAACGAGGAGGCTGGAGAGACTTACCTCTGTGGTCTCAGAATC 480
 DB 722 ATCCCTTAAAGAGTCAACGAGGAGGCTGGAGAGACTTACCTCTGTGGTCTCAGAATC 781
 QY 481 ATGATGTACATCTTCTGTGCTTCTCCTCACTGCTGCTGCTCATCGAGATGATATATGC 540

```
Db 782 ATGATGTACATCTCTCTGGTCTTCTCCTCACCCTTGCTGCTCATCGAGATGATATATGC 841
QY 541 TACAGAAAGGTCTCAAAAGCCGAGAGGAGCCCAAGAAACCGCTCTGACTA 593
Db 842 TACAGAAAGGTCTCAAAAGCCGAGAGGAGCCCAAGAAACCGCTAAGTCCA 894

RESULT 7
ID AAH98320 standard; cDNA; 978 BP.
XX AC AAH98320;
XX DT 12-OCT-2001 (first entry)
XX DE Human EST-derived coding sequence SEQ ID NO: 177.
XX KW Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
XX KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
XX KW diagnostics; forensic test; gene mapping; genetic disorder; biodiversity;
XX KW gene therapy; nutrition; ss.
XX OS Homo sapiens.
XX PN WO200154477-A2.
XX PD 02-AUG-2001.
XX PF 25-JAN-2001; 2001WO-US002687.
XX XX 25-JAN-2000; 2000US-00491404.
XX PR 17-JUL-2000; 2000US-00617746.
XX PR 03-AUG-2000; 2000US-00631451.
XX PR 15-SEP-2000; 2000US-00663870.
XX XX (HYSE-) HYSEQ INC.
XX PA Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
XX PI Cao Y, Drmanac RA, Zhang J, Wehrman T;
XX DR WPI: 2001-476164/51.
XX DR P-PSDB; ANM23661.
XX PT Isolated polypeptide for treatment of diseases, diagnostics, raising
XX PT antibodies and research use.
XX PS Claim 1; Page 299-300; 1275pp; English.
XX CC The present invention provides the protein and coding sequences of novel
XX CC proteins from a variety of organisms, including human, dog, cat, horse,
XX CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
XX CC urchin and tomato. These were derived from expressed sequence tags (ESTs)
XX CC from the organism of interest. They can be used in diagnostics,
XX CC forensics, gene mapping, identification of mutations, to assess
XX CC biodiversity and for nutritional purposes. The present sequence is a cDNA
XX CC of the invention
XX SQ Sequence 978 BP; 193 A; 283 C; 293 G; 209 T; 0 U; 0 Other;

Query Match 88.0%; Score 570.2; DB 4; Length 978;
Best Local Similarity 94.8%; Pred. No. 4.2e-144;
Matches 612; Conservative 0; Mismatches 33; Indels 2; Gaps 2;

QY 1 ATGCCCTGCCTTCAATAGATTGTTCCCTGGCTTCTCTGCTGCTATCTACTGGGTCACT 60
Db 295 ATGCCCTGCCTTCAATAGATTGTTCCCTGGCTTCTCTGCTGCTATCTACTGGGTCACT 354
QY 61 GTCTGCTTCCCTGTGTGTGGAAAGTCCCTCGAGACGAGCCCTGAGGGCAACCCC 120
Db 355 GTCTGCTTCCCTGTGTGTGGAAAGTCCCTCGAGACGAGCCCTGAGGGCAACCCC 414
QY 121 ATGAAGCTGGCTGCATCTCTCTCATGAAGAGAGAGGTGGAGGCCACCGGTGGTG 180
```

```
Db 415 ATGAAGCTGGCTGCATCTCTCTCATGAAGAGAGAGGTGGAGGCCACCGGTGGTG 474
QY 181 GAATGTTCTACAGGCCCGAGGGCGGTAAAGATTTCCTTTATTACAGTATCGGAATGC 240
Db 475 GAATGTTCTACAGGCCCGAGGGCGGTAAAGATTTCCTTTATTACAGTATCGGAATGC 534
QY 241 CACGAGGAGGTGGAGAGCCCTTTTCAGGGGCCCTGCTGAGTGAATGGCAGAGGACCTG 300
Db 535 CACGAGGAGGTGGAGAGCCCTTTTCAGGGGCCCTGCTGAGTGAATGGCAGAGGACCTG 594
QY 301 CAGGAGCTGTCCATCACTGTGCTCAACGTCACCTCTGAACGACTCTCGGCTCTACACCTGC 360
Db 595 CAGGAGCTGTCCATCACTGTGCTCAACGTCACCTCTGAACGACTCTCGGCTCTACACCTGC 654
QY 361 AATGTGTCCCGGAGTTTGAGTTTGAGGCGCATCGGCCCTTTGTGAAGACGACGCGGCTG 420
Db 655 AATGTGTCCCGGAGTTTGAGTTTGAGGCGCATCGGCCCTTTGTGAAGACGACGCGGCTG 714
QY 421 ATCCCTCTTAAGAGTACCGAGAGGCTGGAGAGGACTTCACCTCTGTGTCTCAGAAATC 480
Db 715 ATCCCTCTTAAGAGTACCGAGAGGCTGGAGAGGACTTCACCTCTGTGTCTCAGAAATC 774
QY 481 ATGATGTACATCTCTCTGTGCTCTTCTCACCTGTGTGCTCTCATCGAGATGATATATG- 539
Db 775 ATGATGTACATCTCTCTGTGCTCTTCTCACCTGTGTGCTCTCATCGAGATGATATATG- 834
QY 540 CTACAGAAAGGTCTCAAAAGCCGAGAGGAGCCCAAGAAACCGGTCTTGACTACC-TTG 598
Db 835 CTACAGAAAGGTCTCAAAAGCCGAGAGGAGCCCAAGAAACCGGTCTTGACTACC-TTG 894
QY 599 CCATCCCTCTCAGAAACGAGAGGAGTCTGCGGTACCTGAGGAGGAA 645
Db 895 CGATTTCATTTGAGAAACGAGGAGGAAATTTCTCGGTACCTGCGGGGGA 941

RESULT 8
ADSL11487
ID ADSL11487 standard; DNA; 978 BP.
XX AC ADSL11487;
XX DT 16-DEC-2004 (first entry)
XX DE Human therapeutic contig DNA - SEQ ID 1724.
XX KW antiinflammatory; neuroprotective; antianaemic; cytostatic; vulnery;
XX KW inflammatory; haematopoiesis; immunity; neurodegenerative; stem cell;
XX KW aplastic anaemia; cancer; wound healing; gene therapy; ds; gene.
XX OS Homo sapiens.
XX PN WO2004080148-A2.
XX PD 23-SEP-2004.
XX PF 30-SEP-2003; 2003WO-US030720.
XX PR 02-OCT-2002; 2002US-0416186P.
XX XX (NUVE-) NUVELO INC.
XX PA Tang YT, Asundi V, Ren F, Zhang J, Wehrman T, Wang Z, Ma Y;
XX PI Wang D, Chen R, Zhao QA, Wang J, Ghosh M, Xue AJ, Weng G, Zhou P;
XX DR WPI: 2004-668857/65.
XX DR P-PSDB; ADS12085.
XX PT New polynucleotide, useful in preparing a composition for diagnosing or
XX PT treating inflammatory, neurodegenerative or stem cell disorders, e.g.,
XX PT aplastic anemia or cancer for promoting wound healing.
XX PS Example 2; SEQ ID NO 1724; 718pp; English.
```


XX The invention relates to a novel isolated polynucleotide and the encoded
CC polypeptide. The molecules of the invention demonstrate antiinflammatory,
CC neuroprotective, antianaemic, cytoskeletal and vulnerary activities and may
CC be useful in preparing a composition for diagnosing or treating
CC inflammatory, haematopoietic, immune, neurodegenerative or stem cell
CC disorders, such as aplastic anaemia or cancer, as well as for promoting
CC wound healing. The molecules may also be utilised during gene therapy
CC procedures. The current sequence is that of a human therapeutic contig
CC DNA of the invention. The current sequence is not shown explicitly within
CC the specification but can be accessed from the WIPO web-site.
XX

SQ Sequence 978 BP; 193 A; 283 C; 293 G; 209 T; 0 U; 0 Other;

Query Match 88.0%; Score 570.2; DB 13; Length 978;
Best Local Similarity 94.6%; Pred. No. 4.2e-144;
Matches 612; Conservative 0; Mismatches 33; Indels 2; Gaps 2;

QY 1 ATGCTGCTTCAATAGATTGTTTCCCTGGCTTCTCTCGTCTTATCTACTGGTCACT 60
DB |||||
QY 61 GTCTGCTTCCCTGTGTGTGGAAGTGCCTCGAGAGCGGCGGTGCGAGCAACCCC 120
DB |||||
QY 355 GTCTGCTTCCCTGTGTGTGGAAGTGCCTCGAGAGCGGCGGTGCGAGCAACCCC 414
DB |||||
QY 121 ATGAAGCTGCGCTGCATCTCTGATGAAGAGAGAGAGGTGAGGCCACCCAGTGTG 180
DB |||||
QY 415 ATGAAGCTGCGCTGCATCTCTGATGAAGAGAGAGAGGTGAGGCCACCCAGTGTG 474
DB |||||
QY 181 GAATGTTTCTACAGCCCGGAGCGGTAAAGATTTCCTTATTACGAGTATCGGAATGGC 240
DB |||||
QY 475 GAATGTTTCTACAGCCCGGAGCGGTAAAGATTTCCTTATTACGAGTATCGGAATGGC 534
DB |||||
QY 241 CACAGAGGTGAGAGCCCTTTTCAGGGCGCTGCAAGTGAATGAGCAAGCACTG 300
DB |||||
QY 535 CACAGAGGTGAGAGCCCTTTTCAGGGCGCTGCAAGTGAATGAGCAAGCACTG 594
DB |||||
QY 301 CAGAGCTGTCCATCACTGTGCTCAAGTCACTGCAAGCACTGTGCGCTTACACCTGC 360
DB |||||
QY 595 CAGAGCTGTCCATCACTGTGCTCAAGTCACTGCAAGCACTGTGCGCTTACACCTGC 654
DB |||||
QY 361 AATGTGTCGCGGAGTTTGAAGTTTGAAGCGCATCGGCCCTTTTGAAGAGCAAGCGGCTG 420
DB |||||
QY 655 AATGTGTCGCGGAGTTTGAAGTTTGAAGCGCATCGGCCCTTTTGAAGAGCAAGCGGCTG 714
DB |||||
QY 421 ATCCCCCTAAGATCAAGAGAGGCTGAGAGACTTTCACCTCTGTGCTTCAAGAAATC 480
DB |||||
QY 715 ATCCCCCTAAGAGTCAAGAGAGGCTGAGAGACTTTCACCTCTGTGCTTCAAGAAATC 774
DB |||||
QY 481 ATGATGTACATCTTCTGTGCTTCTCCTCAGCTGTGCTGCTCATCGAGATGATATTTG- 539
DB |||||
QY 775 ATGATGTACATCTTCTGTGCTTCTCCTCAGCTGTGCTGCTCATCGAGATGATATTTG- 834
DB |||||
QY 540 CTACAGAAAGGTCTCAAAAGCCGAGAGCGAGCCCAAGAAAACCGCTCTGACTACC- 598
DB |||||
QY 835 CTACAGAAAGGTCTCAAGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG- 894
DB |||||
QY 599 CCATCCCATCTGAGAACAGAGAACTCTGCGGTACCAAGTGGAGGAA 645
DB |||||
QY 895 CGATTCCATTGAGAACAGAGAAATTTCTTGGTACCTCGGGGGGA 941
DB |||||

RESULT 9

AAC67836

ID AAC67836 standard; cDNA; 2220 BP.

XX

AC

XX

XX

XX

XX

XX

XX

KW Rat; beta sub-unit; beta3; analgesic; anticonvulsant; cerebroprotective;
KW vasotrophic; cardiant; nootropic; cytoskeletal; dermatological;
KW gene therapy; voltage-gated sodium channel; pain; epilepsy; stroke;
KW ischaemia; heart disease; Jacobsen Syndrome;
KW familial nonchromaffin paraganglioma; phenylketonuria;
KW Charcot Marie Tooth disease; ss.
XX

OS Rattus sp.

XX WO200063367-A1.

XX PD 26-OCT-2000.

XX PF 24-FEB-2000; 2000WO-EP001783.

XX PR 15-APR-1999; 99US-0129473P.

XX PA (WARN) WARNER LAMBERT CO.

XX PA (UYCA-) UNIV CAMBRIDGE TECH SERVICES LTD.

XX PI Cox P, Dixon A, Jackson A, Morgan K;

XX DR WPI; 2000-665241/64.

XX DR P-PSDB; AAB36001.

PT Novel nucleic acids encoding a beta-3 subunit from a voltage-gated sodium
PT channel, and their corresponding polypeptides, useful for detecting and
PT treating sodium channel-associated conditions, e.g. pain, epilepsy and
PT stroke.
XX

PS Claim 6; Page 69-70; 88pp; English.

CC The present sequence is given in the claims of a specification relating
CC to a novel family of beta sub-unit proteins from a voltage-gated sodium
CC channel. Human and rat beta sub-units, which have been collectively
CC identified as beta3, have been isolated. The polynucleotides and
CC polypeptides are useful for screening for agonists and antagonists of
CC sodium channels. The agonists, antagonists, proteins and nucleic acids
CC may be used diagnosing of treating diseases or conditions associated with
CC voltage-gated sodium channels, e.g. pain, epilepsy, stroke, ischaemia,
CC heart disease, Jacobsen Syndrome, Familial Nonchromaffin Paraganglioma,
CC Phenylketonuria and Charcot Marie Tooth disease
XX

SQ Sequence 2220 BP; 573 A; 557 C; 561 G; 529 T; 0 U; 0 Other;

Query Match 82.8%; Score 536.6; DB 3; Length 2220;

Best Local Similarity 89.3%; Pred. No. 7.2e-135;

Matches 578; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 1 ATGCTGCTTCAATAGATTGTTTCCCTGGCTTCTCTCGTCTTATCTACTGGTCACT 60
DB |||||
QY 363 ATGCTGCTTCAATAGATTGTTTCCCTGGCTTCTCTCGTCTTATCTACTGGTCACT 422
DB |||||
QY 61 GTCTGCTTCCCTGTGTGTGGAAGTGCCTCGAGAGCGGCGGTGCGAGCAACCCC 120
DB |||||
QY 423 GTCTGCTTCCCTGTGTGTGGAAGTGCCTCGAGAGCGGCGGTGCGAGCAACCCC 482
DB |||||
QY 121 ATGAAGCTGCGCTGCATCTCTCGATGAAGAGAGAGGTGAGGCCACCCAGTGTG 180
DB |||||
QY 483 ATGAAGCTGAGTGCATCTCTCGATGAAGAGAGAGGTGAGGCCACCCAGTGTG 542
DB |||||
QY 181 GAATGTTTCTACAGGCCCGGCGGTAAAGATTTCCTTATTACGATATCGGAATGGC 240
DB |||||
QY 543 GATGTTTCTACAGGCCCGGCGGTAAAGATTTCCTTATTACGATATCGGAATGGC 602
DB |||||
QY 241 CACAGAGGTGAGAGCCCTTTTCAGGGCGGCTGCAAGTGGATGCGAGCAAGCACTG 300
DB |||||
QY 603 CACAGAGGTGAGAGCCCTTTTCAGGGCGGCTGCAAGTGGATGCGAGCAAGCACTG 662
DB |||||
QY 301 CAGAGGTGCTCATCTGCTGCTCAAGCTCACTCTGAACTCTGCGCTTACACCTGC 360
DB |||||
QY 663 CAGAGGTATCCATCACTGCTACTCAATGTCACTTTGAATGACTCTGCGCTTACACATGC 722
DB |||||

QY 361 AATGTGTCGCGGAGTTTCAGTTTGTAGGCGCATCGGCCCTTTGTGAAGACGACGCGCTG 420
 |||||
 Db 723 AATGTGTCAGGAGTTTCGAATTCGAGGACACAGGCCCTTTGTGAAGACCGAGACTG 782
 |||||
 QY 421 ATCCCTTAAGAGTCAACGAGAGGCTGGAGAGGACTTCACCTCTCTGTGTCTCAGAAATC 480
 |||||
 Db 783 ATACCTTTGCGAGTCACTGAAGAGCGGAGAGACTTCACCTCCCTGTCGCTCGGAAATC 842
 |||||
 QY 481 ATGATGTACATCTCTGCTCTTCCTCAACCTGTGCTGCTCATCGAGATGATATATTCG 540
 |||||
 Db 843 ATGATGTACATCTCTGCTCTTCCTCAACCTGTGCTGCTTTATTCAGATGATCTATTCG 902
 |||||
 QY 541 TACAGAAAGTCTCAAAAGCCGAAGAGGAGCGCCCAAGAAACGCGTCTGACTACCTTTCG 600
 |||||
 Db 903 TACAGAAAGTCTCTAAGGCCGGAAGAGGAGGACACAGGAAATCGCTGACTACCTTTCG 962
 |||||
 QY 601 ATCCCTTCAGAAACGAGGAGAACTCTGCGGTACCACTGAGGAGAAATA 647
 |||||
 Db 963 ATCCCTTCAGAAACGAGGAGAACTCTGCGGTACCTGTGGAGAAATA 1009
 |||||
 RESULT 10
 AAC90602
 ID AAC90602 standard; DNA; 2632 BP.
 AC AAC90602;
 XX
 XX
 DT 13-MAR-2001 (first entry)
 DE
 DE Rat sodium channel beta3 protein Alrxa94h5 related sequence.
 XX
 XX Rat; sodium channel beta3 protein; Alrxa94h5; pain; sleep disorder;
 KW neurodegenerative disorder; mood disorder; muscle contraction; ds.
 XX
 OS Rattus sp.
 XX
 XX WO200069912-A1.
 XX
 XX 23-NOV-2000.
 XX
 PF 12-MAY-2000; 2000WO-US013144.
 XX
 PR 14-MAY-1999; 99US-0134198P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Curtis RAJ;
 XX
 XX WPI; 2001-122743/13.
 DR
 XX New rat sodium channel beta-3 subunit gene isolated from a rat dorsal
 PT root ganglion cDNA library for use in chromosome mapping, forensic
 PT medicine, monitoring clinical trials and therapeutics.
 XX
 PS Claim 1; Page 41-42; 145pp; English.
 XX
 CC The present invention provides the protein and coding sequences of the
 CC rat sodium channel beta3 protein, designated Alrxa94h5. This protein is
 CC involved in the generation of pain and other sensory or perceptible nerve
 CC impulses, in the establishment and endurance of mood, neurodegenerative
 CC and sleep disorders, and in the control of muscle contraction, including
 CC movements such as the heartbeat, digestion and vascular tone. The
 CC sequences can be used in predictive medicine, screening and diagnostic
 CC assays, and in pharmacogenomics
 XX
 SQ Sequence 2632 BP; 693 A; 636 C; 611 G; 692 T; 0 U; 0 Other;
 Query Match 82.8%; Score 536.6; DB 5; Length 2632;
 Best Local Similarity 89.3%; Pred. No. 7.6e-135;
 Matches 578; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
 1 ATGCTGCTCAATAGATGTTTCCCTGGCTCTCTGCTGCTATCTACTGGGTCACT 60
 |||||
 QY

Db 78 ATGCTGCTTCAACAGATTGCTTCCCTAGCTTCTCTAGTGTCTCATCTACTGGGTGAGA 137
 |||||
 QY 61 GTCTGCTTCCCTGTGTGTGTGGAAGTGCCTCGGAGACGAGAGCCCTGAGGGCAACCCC 120
 |||||
 Db 138 GTCTGCTTCCCTGTGTGTGTGGAAGTGCCTCGGAGACAGAAAGCGGTGAGGGCAATCCC 197
 |||||
 QY 121 ATGAAGCTGCGCTGCATCTCTGCAATGAAGAGAGGAGGTGGAGGCCACACGCGTGGTG 180
 |||||
 Db 198 ATGAAGCTGAGGTGCATCTCTGCAATGAAGAGGAGAGGTGGAGGCCACCACTGTGGTG 257
 |||||
 QY 181 GAATGTTCTACAGGCCCGAGGGCGGTAAAGATTTCTTTATTACAGTATCGGAATGGC 240
 |||||
 Db 258 GAGTGGTTCTACAGGCCCTGAGGGCGGTAAAGATTTCTTTATATGAGTATCGGAATGGC 317
 |||||
 QY 241 CACAGAGAGGTGGAGAGCCCTTTTCAGGGCGCTGTCAGTGAATGGCAGCAAGGACCTG 300
 |||||
 Db 318 CACAGAGAGGTGGAGAGCCCTTTCCAAGGCCGTCTGCAATGGAATGGAGCAAGACCTG 377
 |||||
 QY 301 CAGGAGGTGTCATCACTGTGCTCAACGTCACCTCTGAACGACTCTCGGCTCTACACCTGC 360
 |||||
 Db 378 CAGGAGGTATCCATCACTGTACTCAATGTCACTTTGAATGACTCTGGCTCTACACATGC 437
 |||||
 QY 361 AATGTGTCGCGGAGTTTTCAGTTTGTAGGCGCATCGGCCCTTTGTGAAGACGACGCGCTG 420
 |||||
 Db 438 AATGTGTCAGGAGTTTCGAATTCGAGGACACAGGCCCTTTGTGAAGACCGAGACTG 497
 |||||
 QY 421 ATCCCTTAAGAGTCAACGAGAGGCTGGAGAGGACTTCACCTCTGTGTCTCAGAAATC 480
 |||||
 Db 498 ATACCTTTGCGAGTCACTGAAGAGCGGAGAGAACTTCACCTCGGTGCTCGGAAATC 557
 |||||
 QY 481 ATGATGTACATCTCTGCTGCTTCTCTCACTGCTGTGCTGCTCATCGAGATGATATATTC 540
 |||||
 Db 558 ATGATGTACATCTCTGCTGCTTCTCTCACTGCTGTGCTGCTTTATTGAGATGATCTATTC 617
 |||||
 QY 541 TACAGAAAGTCTCAAAAGCCGAAGAGGAGCGGCCCAAGAAACGCGTCTCACTACCTTTCG 600
 |||||
 Db 618 TACAGAAAGTCTCTAAGGCCGGAAGAGGAGGACACAGGAAATCGGTCTGACTACCTTGT 677
 |||||
 QY 601 ATCCCTTCAGAAACGAGGAGAACTCTGCGGTACCACTGAGGAGAAATA 647
 |||||
 Db 678 ATCCCTTCAGAAACGAGGAGAACTCTGCGGTACCTGTGGAGAAATA 724
 |||||
 RESULT 11
 AAC90600
 ID AAC90600 standard; cDNA; 3108 BP.
 AC AAC90600;
 XX
 XX
 DT 13-MAR-2001 (first entry)
 DE
 DE Rat sodium channel beta3 protein Alrxa94h5 coding sequence.
 XX
 XX Rat; sodium channel beta3 protein; Alrxa94h5; pain; sleep disorder;
 KW neurodegenerative disorder; mood disorder; muscle contraction; ss.
 XX
 OS Rattus sp.
 XX
 XX WO200069912-A1.
 XX
 XX 23-NOV-2000.
 XX
 PF 12-MAY-2000; 2000WO-US013144.
 XX
 PR 14-MAY-1999; 99US-0134198P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Curtis RAJ;
 XX
 XX WPI; 2001-122743/13.
 DR
 DR P-PSDB; AAB50243, AAB50245.
 XX

PT New rat sodium channel beta-3 subunit gene isolated from a rat dorsal
PT root ganglion cDNA library for use in chromosome mapping, forensic
PT medicine, monitoring clinical trials and therapeutics.
XX Claim 1; Fig 1; 145pp; English.
XX
CC The present invention provides the protein and coding sequences of the
CC rat sodium channel beta3 protein, designated Alrx94h5. This protein is
CC involved in the generation of pain and other sensory or perceptive nerve
CC impulses, in the establishment and endurance of mood, neurodegenerative
CC and sleep disorders, and in the control of muscle contraction, including
CC movements such as the heartbeat, digestion and vascular tone. The
CC sequences can be used in predictive medicine, screening and diagnostic
CC assays, and in pharmacogenomics
XX
XX Sequence 3108 BP; 830 A; 748 C; 727 G; 803 T; 0 U; 0 Other;
SQ
Query Match 82.8%; Score 536.6; DB 5; Length 3108;
Best Local Similarity 89.3%; Pred. No. 8.1e-135;
Matches 578; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
QY 1 ATGCTCGCTTCATAGATGTTTCCCTGGCTTCTCTCGTCTTACTACTGGTCAGT 60
Db |||||||
QY 78 ATGCTCGCTTCAACAGATTGCTTCCCTAGCTTCTTAGTGCTCATCTACTGGTCAGT 137
Db |||||||
QY 61 GTCTGCTTCCCTGTGTGTGGAAGTGCCCTCGAGACGAGCGCGTGCAGGGCAACCCC 120
Db |||||||
QY 138 GTCTGCTTCCCTGTGTGTGGAAGTGCCCTCGAGACGAGCGCGTGCAGGGCAATCCC 197
Db |||||||
QY 121 ATGAAGTGGCTGCATCTCTCTCATGAAGAGAGAGAGTGGAGGCCACACCGTGGTG 180
Db |||||||
QY 198 ATGAAGCTGAGTGCATCTCTCTCATGAAGAGAGAGAGTGGAGGCCACACCTGTGGTG 257
Db |||||||
QY 181 GAATGTTCTACAGCGCGGCGGTAAAGATTTCCTTATTACGATATCGAAATGCG 240
Db |||||||
QY 258 GAGTGGTCTACAGCGCTGAGGCGGTAAAGATTTCCTTATATAGATATCGGAATGCG 317
Db |||||||
QY 241 CACGAGAGTGGAGAGCGCCCTTTCAGGGCGCTCGAGTGAATGGCAGCAAGGACCTG 300
Db |||||||
QY 318 CACGAGAGTGGAGAGCGCCCTTTCAGGGCGCTCGAGTGAATGGCAGCAAGGACCTG 377
Db |||||||
QY 301 CAGAGCTGTCCATCATCTGTGCTCAACGTCACTCTGAACGACTCTGCGCTTACACCTGC 360
Db |||||||
QY 378 CAGGAGCTATCCATCACTACTCAATGTCACTTTGAATGACTCTGCGCTTACACATGC 437
Db |||||||
QY 361 AATGTGTCGGGAGTTGAGTTGAGCGGCATCGGCCCTTTTGGAGAGACGCGGCTG 420
Db |||||||
QY 438 AATGTGTCGGGAGTTGGAATTCGAGGCGACACAGGCGCTTTTGTGAAGACCCAGAGCTG 497
Db |||||||
QY 421 ATCCCCCTAAGAGTCAACGAGGAGGCTGGAGAGACTTTCACCTCTGTGGTCTCAGAAATC 480
Db |||||||
QY 498 ATACCTTTGCGAGTCACTGAAGAGCGGGAGAGACTTTCACCTCGTGGTCTCGAAATC 557
Db |||||||
QY 481 ATGATGTACATCTTCTGCTTCTCCTCACCTGTGGTCTGCTCAFCGAGATGATATATGC 540
Db |||||||
QY 558 ATGATGTACATCTTCTGCTTCTCCTCACCTGTGGTCTGCTTATGAGATGATATATGC 617
Db |||||||
QY 541 TACAGAAAGTCTCAAAAGCGGAGAGCGAGCGCCCAAGAAACCGCTCTGACTACCTTGC 600
Db |||||||
QY 618 TACAGAAAGTCTCTAAGGCGGAGAGCGAGCGACAGGAAATGCGTCTGACTACCTTGC 677
Db |||||||
QY 601 ATCCCATCTGAGAAACAGGAGAACTCTGCGGTACACGTCAGTGAGGGAATA 647
Db |||||||
QY 678 ATCCCTTCAGAGAACAGGAGAACTCTGTGGTACCTGTGGAGGAAATA 724
Db |||||||

RESULT 12
AAS86764
ID AAS86764 standard; cDNA; 1195 BP.
XX
AC AAS86764;
XX
DT 13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #22568.
DE
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
PD
XX
XX 30-MAR-2001; 2001WO-US008631.
XX
XX 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
PA Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
DR P-PSDB; ABG22577.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX Claim 1; SEQ ID NO 22568; 103pp; English.
PS
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 1195 BP; 253 A; 356 C; 381 G; 205 T; 0 U; 0 Other;
SQ
Query Match 82.5%; Score 534.8; DB 5; Length 1195;
Best Local Similarity 95.7%; Pred. No. 1.8e-134;
Matches 572; Conservative 0; Mismatches 22; Indels 4; Gaps 2;
QY 1 ATGCTCGCTTCAATAGATTGTTTCCCTGGCTTCTCTCGTCTTACTACTGGTCAGT 60
Db 463 ATGCTCGCTTCAATAGATTGTTTCCCTGGCTTCTCTCGTCTTACTACTAGTCACT 522
QY 61 GTCTGCTTCCCTGTGTGTGGAA--GTGCGCTCGAGACGAGCGCGTGA--GGGCAA 116
Db 523 GTCTGCTTCCCTGTGTGTGGGGAAGTGCCCTTAGAAAAACGGGGGGCGGCTAA 582
QY 117 CCCATGAAGCTCGCTGTGATCTCTCATGAAGAGAGAGAGTGGAGGCCACACCGT 176
Db 583 CCCATGAAGCTCGCTGTGATCTCTCATGAAGAGAGAGAGTGGAGGCCACACCGT 642
QY 177 GGTGGAATGTTTCTACAGGCCCGGAGCGGCGTAAAGATTTCCTTATTTACGAGTATCGGAA 236
Db |||||||

Db 643 GGTGGAATGGTTCTACAGGCCCGGCGGTAAGATTTCTTATTTACAGATATCGGAA 702
 QY 237 TGGCCACCAAGGAGTGGAGAGCCCTTTTTCAGGGGCGCCTGCAAGTGGAAATGGCAGCAAGGA 296
 Db 703 TGGCCACCAAGGAGTGGAGAGCCCTTTTTCAGGGGCGCCTGCAAGTGGAAATGGCAGCAAGGA 762
 QY 297 CTTGAGGAGAGTGTCCATCACTGTGTCTCAACGTCACTGTGAACGACTCTGGGCTCTACAC 356
 Db 763 CTTGAGGAGAGTGTCCATCACTGTGTCTCAACGTCACTGTGAACGACTCTGGGCTCTACAC 822
 QY 357 CTGCAATGTCTCCGGGAGTTTGAAGTTTTCAGGGCGCATCGGCCCTTTTGAAGACGACGCG 416
 Db 823 CTGCAATGTCTCCGGGAGTTTGAAGTTTTCAGGGCGCATCGGCCCTTTTGAAGACGACGCG 882
 QY 417 GCTGATCCCCCTAAGAGTCAACGAGGAGGCTGGAGAGGACTTCACTCTGTGTCTCAGA 476
 Db 883 GCTGATCCCCCTAAGAGTCAACGAGGAGGCTGGAGAGGACTTCACTCTGTGTCTCAGA 942
 QY 477 AATCATGATGTACATCTTCTGTGTCTTCTCCTCACCCTGTGGCTCTCATCGAGATGATATA 536
 Db 943 AATCATGATGTACATCTTCTGTGTCTTCTCCTCACCCTGTGGCTCTCATCGAGATGATATA 1002
 QY 537 TTGCTACAGAAAGGTCTCAAAAGCCGAAGAGGCGAGCCCAAGAAACGCGTCTGACTAC 594
 Db 1003 TTGCTACAGACAGGTCTCAAAAGCCGAAGAGGCGAGCCCAAGAAACGCGCCAGGAAC 1060

RESULT 13

AAC90601
 ID AAC90601 standard; cDNA; 645 BP.
 XX
 AC AAC90601;
 XX
 DT 13-MAR-2001 (first entry)
 XX
 DE Rat sodium channel beta3 protein Alrxa94h5 partial coding sequence.
 XX
 KW Rat; sodium channel beta3 protein; Alrxa94h5; pain; sleep disorder;
 KW neurodegenerative disorder; mood disorder; muscle contraction; ss.
 XX
 OS Rattus sp.
 XX
 PN WO200069912-A1.
 XX
 PD 23-NOV-2000.
 XX
 PF 12-MAY-2000; 2000WO-US013144.
 XX
 PR 14-MAY-1999; 99US-0134198P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Curtis RAJ;
 XX
 DR WPI; 2001-122743/13.
 XX
 PT New rat sodium channel beta-3 subunit gene isolated from a rat dorsal
 PT root ganglion cDNA library for use in chromosome mapping, forensic
 PT medicine, monitoring clinical trials and therapeutics.
 XX
 PS Claim 1; Page 40-41; 145pp; English.
 XX

CC The present invention provides the protein and coding sequences of the
 CC rat sodium channel beta3 protein, designated Alrxa94h5. This protein is
 CC involved in the generation of pain and other sensory or perceptible nerve
 CC impulses, in the establishment and endurance of mood, neurodegenerative
 CC and sleep disorders, and in the control of muscle contraction, including
 CC movements such as the heartbeat, digestion and vascular tone. The
 CC sequences can be used in predictive medicine, screening and diagnostic
 CC assays, and in pharmacogenomics
 XX
 SQ Sequence 645 BP; 155 A; 154 C; 181 G; 155 T; 0 U; 0 Other;

Query Match 82.5%; Score 534.6; DB 5; Length 645;
 Best Local Similarity 89.3%; Pred. No. 1.6e-134;
 Matches 576; Conservative 0; Mismatches 69; Indels 0; Gaps 0;
 QY 1 ATGCGTCCCTTCAATAGATTGTTTCCCTGGGCTTCTCTCGTGGTATTCTACTTGGGTCACT 60
 Db 1 ATGCGTCCCTTCAACAGATTGTTTCCCTTAGCTTCTCTAGTGTCTCATCTACTGGGTCA 60
 QY 61 GTCTGCTTCCCTGTGTGTGGAAGTCCCTCGGAGACGAGGCGGTGAGGGCAACCC 120
 Db 61 GTCTGCTTCCCTGTGTGTGGAAGTCCCTCGGAGACAGAAAGCGGTGAGGGCAATCC 120
 QY 121 ATGAAGCTCGCTGCTATCTCTGATGAGAGAGGAGGTGAGGCCACCACTGTGGTG 180
 Db 121 ATGAAGCTGAGGTGCTATCTCTGATGAGAGGAGGAGGTGAGGCCACCACTGTGGTG 180
 QY 181 GAATGTTCTACAGGCCCGGAGGCGGTAAAGATTTCCTTATTTACGAGTATCGGAATGGC 240
 Db 181 GAGTGGTCTACAGGCCCTGAGGCGGTAAAGATTTCCTTATATAGATATCGGAATGGC 240
 QY 241 CACGAGGAGGTGAGAGGCCCTTTTCAGGGGCGCTGCACTGGAATGGCAGCAAGCACTG 300
 Db 241 CACGAGGAGGTGAGAGGCCCTTTTCCAAAGCGCTCTGCACTGGAATGGGAGCAAGCACTG 300
 QY 301 CAGGAGGTGCTCATCACTGTGTCTCAACGTCACTCTGAACGACTCTGGCCTCTACACCTGC 360
 Db 301 CAGGAGGTATCCATCACTGTGTCTCAATGTCACTTTGAATGACTCTGGCCTCTACACATGC 360
 QY 361 AATGTGTCCCGGAGTTTTCAGTTTTCAGGCGCATCGGCCCTTTTGTGAAGACGCGGCTG 420
 Db 361 AATGTGTCCAGGAGTTTCGAATTCGAGGCCACACAGGCCCTTTTGTGAAGACGAGACTG 420
 QY 421 ATCCCTTAAAGTCAACGAGGAGGTGAGAGGACTTCACTCTGTGTGTCTCAGAAATC 480
 Db 421 ATACCTTTCCGAGTCACTGAAGAGCGGAGAGACTTCACTCTGTGTGTCTCAGAAATC 480
 QY 481 ATGATGTACATCTTCTGTGTCTTCTCCTCACCCTGTGGTCTCTCATCGAGATATATTGC 540
 Db 481 ATGATGTACATCTTCTGTGTCTTCTCCTCACCCTGTGGTCTTATTTAGATGATCTATTGC 540
 QY 541 TACAGAAAGGTCTCAAAAGCCGAAGAGGCGAGCCCAAGAAACCGGTCTGACTACCTTGC 600
 Db 541 TACAGAAAGGTCTCTAAAGCCGAAGAGGCGAGCCCAAGAAATGCGTCTGACTACCTTGC 600
 QY 601 ATCCCATCTGAGAACAGGAGAACTCTGCGGTACCACTGAGGAGAA 645
 Db 601 ATCCCTTTCAGAGAACAAAGGAGAACTCTGTGTACCTGTGGAGAA 645

RESULT 14

ADS10151
 ID ADS10151 standard; DNA; 953 BP.
 XX
 AC ADS10151;
 XX
 DT 16-DEC-2004 (first entry)
 XX
 DE Human therapeutic DNA - SEQ ID 388.
 XX
 KW antiinflammatory; neuroprotective; antianaemic; cytostatic; vulnery;
 KW inflammatory; haematopoiesis; immunity; neurodegenerative; stem cell;
 KW aplastic anaemia; cancer; wound healing; gene therapy; ds; gene.
 XX
 OS Homo sapiens.
 XX
 PN WO2004080148-A2.
 XX
 PD 23-SEP-2004.
 XX
 PF 30-SEP-2003; 2003WO-US030720.
 XX
 PR 02-OCT-2002; 2002US-0416186P.
 XX

PA (NUVE-) NUVELO INC.
 XX Tang YT, Asundi V, Ren F, Zhang J, Wehrman T, Wang Z, Ma Y;
 PI Wang D, Chen R, Zhao QA, Wang J, Ghosh M, Xue AJ, Weng G, Zhou P;
 XX WPI; 2004-668857/65.
 DR P-PSDB; ADS10835.
 XX New polynucleotide, useful in preparing a composition for diagnosing or
 PT treating inflammatory, neurodegenerative or stem cell disorders, e.g.,
 PT aplastic anemia or cancer for promoting wound healing.
 XX Claim 1; SEQ ID NO 388; 718pp; English.
 XX The invention relates to a novel isolated polynucleotide and the encoded
 CC polypeptide. The molecules of the invention demonstrate anti-inflammatory,
 CC neuroprotective, antianemic, cytostatic and vulnerary activities and may
 CC be useful in preparing a composition for diagnosing or treating
 CC inflammatory, haematopoietic, immune, neurodegenerative or stem cell
 CC disorders, such as aplastic anaemia or cancer, as well as for promoting
 CC wound healing. The molecules may also be utilised during gene therapy
 CC procedures. The current sequence is that of a human therapeutic DNA of
 CC the invention. The current sequence is not shown explicitly within the
 CC specification but can be accessed from the WIPO web-site.
 XX
 SQ Sequence 953 BP; 180 A; 279 C; 290 G; 204 T; 0 U; 0 Other;
 Query Match 77.8%; Score 504; DB 13; Length 953;
 Best Local Similarity 99.0%; Pred. No. 3.6e-126; Indels 0; Gaps 0;
 Matches 507; Conservative 0; Mismatches 5;
 QY 1 ATGCTGCTCCATAGATGTTTCCCTGGCTTCTCGTGTCTTACTACTGGTCAGT 60
 DB 327 ATGCTGCTCCATAGATGTTTCCCTGGCTTCTCGTGTCTTACTACTGGTCAGT 386
 QY 61 GTCTGCTTCCCTGTGTGGAAGTCCCTCGAGACGAGCGCGGTGAGGGAACCC 120
 DB 387 GTCTGCTTCCCTGTGTGGAAGTCCCTCGAGACGAGCGCGGTGAGGGAACCC 446
 QY 121 ATGAAGCTGCTGCTGCTCTCTCATGAGAGAGAGGAGTGGAGCCACCGGTGGT 180
 DB 447 ATGAAGCTGCTGCTGCTCTCTCATGAGAGAGAGGAGTGGAGCCACCGGTGGT 506
 QY 181 GAATGTTTCTACAGCCGCGGCGGTAAAGATTCTCTATTACGATATCGAATGGC 240
 DB 507 GAATGTTTCTACAGCCGCGGCGGTAAAGATTCTCTATTACGATATCGAATGGC 566
 QY 241 CACGAGAGTGGAGAGCCCTTTACGAGCGCGCTGAGTGGGAATGCGAGAGACCTG 300
 DB 567 CACGAGAGTGGAGAGCCCTTTACGAGCGCGCTGAGTGGGAATGCGAGAGACCTG 626
 QY 301 CAGGAGTGTCCATCACTGCTCAAGTCACTCTGAACGACTCTGCGCTTACACCTGC 360
 DB 627 CAGGAGTGTCCATCACTGCTCAAGTCACTCTGAACGACTCTGCGCTTACACCTGC 686
 QY 361 AATGTGTCCGCGGAGTTTGAAGTGGAGCGCATCGCGCTTTGTGAAGACGACGCGCTG 420
 DB 687 AATGTGTCCGCGGAGTTTGAAGTGGAGCGCATCGCGCTTTGTGAAGACGACGCGCTG 746
 QY 421 ATCCCTTAAAGATCAACGAGAGGCTGGAGAGGATTCACCTCTGTGTGCTCAGAAATC 480
 DB 747 ATCCCTTAAAGATCAACGAGAGGCTGGAGAGGATTCACCTCTGTGTGCTCAGAAATC 806
 QY 481 ATGATGTACATCTTCTGTGTTTCCCTACCCCT 512
 DB 807 ATGATGTACATCTTCTGTGTTTCCCTACCCCT 838
 RESULT 15
 AAV86895
 ID AAV86895 standard; cDNA; 471 BP.
 XX
 AC AAV86895;

XX 27-APR-1999 (first entry)
 XX EST clone BM4.
 DE
 XX Expressed sequence tag; secreted protein; haematopoiesis regulator;
 KW tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;
 KW chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;
 KW receptor; ligand; anti-inflammatory; tumour inhibitor; ds.
 OS Homo sapiens.
 XX WO9845435-A2.
 PN
 XX 15-OCT-1998.
 PD
 XX 10-APR-1998; 98WO-US006954.
 PF
 XX 10-APR-1997; 97US-00835913.
 PR
 XX (GEMY) GENETICS INST INC.
 PA
 XX Jacobs K, McCoy JM, Lavallie ER, Racie LA, Merberg D, Treacy M;
 PI Spaulding V, Agostino MU;
 PI WPI; 1999-070076/06.
 DR
 XX New polynucleotides encoding human secreted proteins - derived from e.g.
 PT human blood, kidney, foetal lung, placenta, testes, brain, ovary,
 PT pituitary, retina and colon cDNA libraries.
 PT
 XX Claim 1; Page 383-384; 633pp; English.
 PS
 XX This sequence represents an expressed sequence tag (EST), and is a
 CC polynucleotide of the invention. The polynucleotides of the invention are
 CC all secreted EST sequences isolated from a variety of human tissue
 CC sources. The EST sequences and proteins encoded by them are predicted to
 CC have useful biological activities which would make them suitable for
 CC treating, preventing or ameliorating medical conditions in humans and
 CC animals, although no supporting data is given. Suggested activities
 CC include nutritional activity, immune stimulating or suppressing activity,
 CC haematopoiesis regulating activity, tissue growth activity,
 CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
 CC activity, cadherin/tumour invasion suppressor activity, tumour inhibition
 CC activity. The EST sequences are also stated to be useful for gene therapy
 CC
 XX
 SQ Sequence 471 BP; 93 A; 126 C; 143 G; 108 T; 0 U; 1 Other;
 Query Match 65.1%; Score 422; DB 2; Length 471;
 Best Local Similarity 99.5%; Pred. No. 4.6e-104;
 Matches 433; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 QY 1 ATGCTGCTTCAATAGATGTTTCCCTGGCTTCTCGTGTCTTACTACTGGTCAGT 60
 DB 29 ATGCTGCTTCAATAGATGTTTCCCTGGCTTCTCGTGTCTTACTACTGGTCAGT 88
 QY 61 GTCTGCTTCCCTGTGTGGAAGTCCCTCGAGACGAGCGCGGTGAGGGAACCC 120
 DB 89 GTCTGCTTCCCTGTGTGGAAGTCCCTCGAGACGAGCGCGGTGAGGGAACCC 147
 QY 121 ATGAAGCTGCTGCTGCTCTCATGAGAGAGAGGAGTGGAGCCACCGGTGGT 180
 DB 148 ATGAAGCTGCTGCTGCTCTCATGAGAGAGAGGAGTGGAGCCACCGGTGGT 207
 QY 181 GAATGTTTCTACAGCCGCGGCGGTAAAGATTCTCTATTACGATATCGAATGGC 240
 DB 208 GAATGTTTCTACAGCCGCGGCGGTAAAGATTCTCTATTACGATATCGAATGGC 267
 QY 241 CACGAGAGTGGAGAGCCCTTTACGAGCGCGCTGAGTGGGAATGCGAGAGACCTG 300
 DB 268 CACGAGAGTGGAGAGCCCTTTACGAGCGCGCTGAGTGGGAATGCGAGAGACCTG 327

QY 301 CAGGACGTGTCATCACTGTGCTCAACGTCACTCACTGTGACGACTCTGGCCTCTACACCTGC 360
 Db |||||
 QY 328 CAGGACGTGTCATCACTGTGCTCAACGTCACTCACTGTGACGACTCTGGCCTCTACACCTGC 387
 Db |||||
 QY 361 AATGTGTCCCGGAGTTTGAGTTTGAGGGCGCATCGGCCCTTTGTGAAGACGACGCGGCTG 420
 Db |||||
 QY 388 AATGTGTCCCGGAGTTTGAGTTTGAGGGCGCATCGGCCCTTTGTGAAGACGACGCGGCTG 447
 Db |||||
 QY 421 ATCCCCCTAAGAGTC 435
 Db |||||
 QY 448 ATCCCCCTAAGAGTC 462
 Db |||||

Search completed: April 7, 2005, 11:09:35
 Job time : 475.712 secs

| Result No. | Score | Query | | Length | DB | ID | Description |
|------------|-------|-------|------|--------|----|---------------------|-------------------|
| | | Match | % | | | | |
| C | 1 | 143.8 | 22.2 | 617 | 4 | US-09-949-016-2369 | Sequence 2369, Ap |
| | 2 | 51 | 7.9 | 11174 | 4 | US-09-949-016-14111 | Sequence 14111, A |
| | 3 | 48.6 | 7.5 | 30337 | 4 | US-09-949-016-13053 | Sequence 13053, A |
| | 4 | 45.4 | 7.0 | 1845 | 4 | US-09-023-587A-5 | Sequence 5, App1 |
| | 5 | 44.8 | 6.9 | 1053 | 4 | US-09-902-540-2866 | Sequence 2666, Ap |
| | 6 | 44.8 | 6.9 | 13706 | 4 | US-09-902-540-1124 | Sequence 1124, Ap |
| | 7 | 42.4 | 6.5 | 2133 | 4 | US-09-902-540-4106 | Sequence 4106, Ap |
| | 8 | 42.4 | 6.5 | 23738 | 4 | US-09-902-540-1203 | Sequence 1203, Ap |
| | 9 | 41.6 | 6.4 | 696 | 4 | US-09-489-039A-3591 | Sequence 3591, Ap |
| | 10 | 40.6 | 6.3 | 2943 | 4 | US-09-902-540-3929 | Sequence 3929, Ap |
| | 11 | 40.6 | 6.3 | 21395 | 4 | US-09-902-540-1394 | Sequence 1194, Ap |
| C | 12 | 40.4 | 6.2 | 1533 | 4 | US-09-902-540-7877 | Sequence 7877, Ap |
| | 13 | 40.4 | 6.2 | 5228 | 4 | US-09-902-540-789 | Sequence 789, App |
| | 14 | 39.2 | 6.0 | 732 | 4 | US-09-902-540-3806 | Sequence 3806, Ap |
| | 15 | 39.2 | 6.0 | 6202 | 4 | US-09-774-528-120 | Sequence 120, App |
| | 16 | 39.2 | 6.0 | 18551 | 4 | US-09-902-540-1187 | Sequence 1187, Ap |
| | 17 | 38.6 | 6.0 | 513 | 4 | US-09-902-540-8812 | Sequence 8812, Ap |
| | 18 | 38.6 | 6.0 | 1407 | 4 | US-09-902-540-6215 | Sequence 6215, Ap |
| | 19 | 38.6 | 6.0 | 8122 | 4 | US-09-902-540-937 | Sequence 937, App |
| | 20 | 38.6 | 6.0 | 13332 | 4 | US-09-902-540-1047 | Sequence 1047, Ap |
| | 21 | 38.2 | 5.9 | 1437 | 4 | US-09-902-540-8578 | Sequence 8578, Ap |
| | 22 | 38.2 | 5.9 | 9081 | 4 | US-09-902-540-905 | Sequence 905, App |
| C | 23 | 37.4 | 5.8 | 711 | 4 | US-09-902-540-6903 | Sequence 6903, Ap |
| | 24 | 37.4 | 5.8 | 4169 | 4 | US-09-902-540-590 | Sequence 590, App |
| | 25 | 37 | 5.7 | 597 | 4 | US-09-902-540-3561 | Sequence 3561, Ap |
| | 26 | 37 | 5.7 | 19019 | 4 | US-09-902-540-1171 | Sequence 1171, Ap |
| | 27 | 36.8 | 5.7 | 6367 | 4 | US-09-902-540-5649 | Sequence 5649, Ap |
| | 28 | 36.8 | 5.7 | | | | |

```
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13053
; LENGTH: 30337
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(30337)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13053

Query Match
Best Local Similarity 7.5%; Score 48.6; DB 4; Length 30337;
Matches 75; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 454 GACTTCACCTCTGCTGCTCAGAAATCATGATGATACATCTTCTGCTTCTCCTCACCCTG 513
Db 620 GACATGGCATCCATCGTCTGAGATCATGATGATGCTCATTTGCTGTTGACCATG 679

QY 514 TGGCTGCTCATCGAGATGATATTTCTACAGAAAGGTTCTCAAAAGCCGGAAGCAGC 572
Db 680 TGGCTGCTGAGAGATGATTTACTGCTACAGAAAGATCGCTGCCGCCACGAGACTGC 738

RESULT 4
US-09-023-587A-5
; Sequence 5, Application US/09023587A
; Patent No. 6653530
; GENERAL INFORMATION:
; APPLICANT: Shewmaker, Christine K.
; APPLICANT: Bhat, Ganesh B.
; APPLICANT: Venkatramesh, Mylavaram
; APPLICANT: Rangwala, Shaikat H.
; APPLICANT: Kishore, Ganesh M.
; APPLICANT: Boddupalli, Sekhar S.
; TITLE OF INVENTION: Methods for Producing Carotenoid Compounds, Tocopherol Compounds,
; TITLE OF INVENTION: Specialty Oils in Plant Seeds
; FILE REFERENCE: 16516.122
; CURRENT APPLICATION NUMBER: US/09/023,587A
; CURRENT FILING DATE: 1998-02-13
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1645
; TYPE: DNA
; ORGANISM: Zea mays
US-09-023-587A-5

Query Match
Best Local Similarity 7.0%; Score 45.4; DB 4; Length 1645;
Matches 151; Conservative 0; Mismatches 176; Indels 0; Gaps 0;

QY 54 GGTCAAGTGTGCTTCCCTGTGTGTGTGTAAGTGCCTCGGAGACGAGGCGGTGACGGG 113
Db 349 GGTCTCAACGGCTCTTCTTAAGGTACGAGCGCCCAAGAGCGGTAACGGTCTGACGT 408

QY 114 CAACCCCATGAGTGGCTGCTCATCTCTGATGAGAGAGGAGTGGAGCCACCCAC 173
Db 409 GGTGCACTACACCACTAGCAGCGGAGCAACCGCAAGGTTCGGCGGAGAACGGTCTGTT 468

QY 174 GGTGTGGAATGTTTCTTACAGCGCCCGCGGTAAAGATTTCTTTATTACGATATCG 233
Db 174 GGTGTGGAATGTTTCTTACAGCGCCCGCGGTAAAGATTTCTTTATTACGATATCG 233
```

```
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14111
; LENGTH: 11174
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(11174)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14111

Query Match
Best Local Similarity 7.9%; Score 51; DB 4; Length 11174;
Matches 90; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 66 CTTCCTGCTGTGTGTAAGTGCCTTCGAGACGAGGCGGTGCGAGCAACCCCATGAA 125
Db 2011 CTGCGGGGCTGCTGAGGTGATCTCGAGACCGAGGCGGTGATGGATGACCTTCAA 2070

QY 126 GCTGCGCTGCATCTCTCTGATGAAGAGAGAGAGGTGGAGGCCACACCGTGTGGAATG 185
Db 2071 AATTCTTTGATCTCTCTGCAAGCGCGGACGAGACCAACGCTGAGACCTTCACCGAGTG 2130

QY 186 GTTCTACAGCCGAGGCGGTAAAGATTTCTTA 220
Db 2131 GACCTTCGCCCAAGAGGCGGACTGAGAGTTGTGCA 2165

RESULT 3
US-09-949-016-13053
; Sequence 13053, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
```


Db 469 CGAGTGGACCGCATCTGTGGCGCGGACGCGGCAACTCTCGGTGGCCAAACGACATGGG 528
QY 234 GAATGGCCACACGAGGTGGAGAGCCCTTTTCAGGGGCGGCTGCAGTGGAAATGCGAGCAA 293
Db 529 CGCGGGCGACTAGAGTACGCCATCGGCTTCAGAGGCGCGTCAAGATCCCGACGACAA 588
QY 294 GGACCTGCAGGACGTGTCCATCACTGTGTCTCAAGCTCACTCTGAACGACTCTGGCCCTCTA 353
Db 589 GATGGTGTACTACGAGGAGCGCGGAGATGTAGTGGCGGACGAGCTCTCTCCGACTT 648
QY 354 CACTGCAATGTGTCCCGGAGTTGA 380
Db 649 CTACGGCTGGGTGTTCCTCCCAAGTCCGA 675

RESULT 5

US-09-902-540-2666
; Sequence 2666, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 2666
; LENGTH: 1053
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-2666

Query Match 6.9%; Score 44.8; DB 4; Length 1053;
Best Local Similarity 52.1%; Pred. No. 0.0076;
Matches 100; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
QY 61 GTCTGCTTCCCTGTGTGTGGAAATGTCCTCGAGACGAGGCGGTCGAGGCAACCC 120
Db 151 GACTTCTCCACGACCTGTGTGGACCTCGCTCGAAGATGACGACGCGGAGCGCCACTC 210
QY 121 ATGAAGTGGCTGCATCTCTCATGAAGAGAGAGAGGTGGAGCCACACGTTGGT 180
Db 211 ATCCACCGGCTGTGTGGCTTCTTCGCGACGGGTGACAGCATCGTCGGCAACACCTGGT 270
QY 181 GAATGGTTCATACAGGCGCGGTAAGATTTCTTATTTACGAGTATCGAATGGC 240
Db 271 CTGAACCTCTACAAGCACCTGAAGCCCTCGAGGCGGGATGTACCTGTGCGGCCAGCTC 330
QY 241 CACAGAGAGGTG 252
Db 331 TACGAGGAGGCG 342

RESULT 6

US-09-902-540-1124/c
; Sequence 1124, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10

; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1124
; LENGTH: 13706
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-1124

Query Match 6.9%; Score 44.8; DB 4; Length 13706;
Best Local Similarity 52.1%; Pred. No. 0.027;
Matches 100; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 61 GTCTGCTTCCCTGTGTGTGGAAATGTCCTCGAGACGAGGCGGTCGAGGCAACCC 120
Db 4948 GACTTCTCCACGACCTGTGTGGACCTCGCTCGAAGATGACGACGCGGAGCGCCACTC 4889
QY 121 ATGAAGTGGCTGCATCTCTCATGAAGAGAGAGAGGTGGAGCCACACGTTGGT 180
Db 4888 ATCCACCGGCTGTGTGGCTTCTTCGCGACGGGTGACAGCATCGTCGGCAACACCTGGT 4829
QY 181 GAATGGTTCATACAGGCGCGGTAAGATTTCTTATTTACGAGTATCGAATGGC 240
Db 4828 CTGAACCTCTACAAGCACCTGAAGCCCTCGAGGCGGGATGTACCTGTGCGGCCAGCTC 4769
QY 241 CACAGAGAGGTG 252
Db 4768 TACGAGGAGGCG 4757

RESULT 7

US-09-902-540-4106/c
; Sequence 4106, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 4106
; LENGTH: 2133
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-4106

Query Match 6.5%; Score 42.4; DB 4; Length 2133;
Best Local Similarity 51.6%; Pred. No. 0.053;
Matches 97; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 87 GCCTTCGAGACGAGGCGGTGAGGGCAACCCCATGAAGCTGGCTGATCTCTCAT 146
Db 1813 GGCCAGGATGATGGCGGCGCCATCCCGTTTCAAGACCGTGCAGCGCGCTGCAG 1754
QY 147 GAAGAGAGAGAGGTGAGGCCACACCGTGTGGATGTGTTTACAGGCCCGAGGCGG 206
Db 1753 GAAGCGGCGGCGGAGTTCGGACACGCTGGGCGTGGGACCGCTTCCCAAGCCAGTGTGT 1694
QY 207 TAAAGATTTCTTATTTACGAGTATCGAATGTCACAGAGGTGGAGAGCCCTTTCA 266
Db 1693 GGGTGAAGCGCTCCGTACAGGATCTGCGCGGACTCCAGCTCGGACAGCGCGCTGCA 1634
QY 267 GGCGCGCC 274
Db 1633 GCTCCGCC 1626

RESULT 8

US-09-902-540-1203

```
; Sequence 1203, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1203
; LENGTH: 23738
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-1203

Query Match 6.5%; Score 42.4; DB 4; Length 23738;
Best Local Similarity 51.6%; Pred. No. 0.17;
Matches 97; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 87 GCGCTCGAGACGGAGCGCTGCAGGGCAACCCCATGAAGCTGCGTGCATCTCTGCAT 146
DB 4191 GCGCAGGATGATGGCGGGCGCCATCCCGTTTGAAGACGCGTGAAGCGCGCGCTGCAG 4250

QY 147 GAAGAGAGAGGAGTGGAGGCCACACACGCTGGTGAATGGTTCTACAGGCCCGAGGGCGG 206
DB 4251 GAAGCGGGCGGGAGTGGACACGCTGGCGTGGAGCGCTTCCAGCACCACTGTGT 4310

QY 207 TAAAGATTTCCTTTATACAGTATCGGAATCGGCACCAAGAGGTGGAGAGCCCTTTCA 266
DB 4311 GGGTGAAGCGCTCCGCTACGAGGATCTGGCGGACTCCAGCTCGGACAGCGCGGTGCA 4370

QY 267 GGGGGGCC 274
DB 4371 GCTCCGCC 4378

RESULT 9
US-09-489-039A-3591
; Sequence 3591, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 3591
; LENGTH: 696
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-3591

Query Match 6.4%; Score 41.6; DB 4; Length 696;
Best Local Similarity 48.3%; Pred. No. 0.053;
Matches 116; Conservative 0; Mismatches 124; Indels 0; Gaps 0;

QY 83 AAGTGCCCTCGAGACGGAGCGGTGCAGGGCAACCCCATGAAGTGGCGCTGCATCTCT 142
DB 92 AGCGCGGACCTCGCGAGGGGTGCTGGGCAAGCTGCCGAAGCGCGCTATCGCAT 151

QY 143 GCATGAAGAGAGAGGTGGAGGCCACCAACCGTGGTGAATGGTTCTACAGGCCCGGAG 202
DB 152 ACCTGCGGTTCAACGACGTTGACGGTACCCCGCTGGATCAGGGGATTCGGCTGTGTTCC 211
```

```
QY 203 GCGTAAAGATTTCCTTATTACAGTATCGAATGGCCACACAGGAGTGGAGAGCCCT 262
DB 212 CCGGGCCGAACCTCTTTTACCGGGGAAGATGTCTTGTGCTGACGGGCCACAGCGGCCGG 271

QY 263 TTCAGGGGCGCTCGAGTGGATGCGACGACCAAGGACCTGCAGGACGTGTCTCATCTGTGC 322
DB 272 TCATTCTGACCTGCTGCTTAAACGTATTCTGACCTGCGGGCTTGCATCGCAGGC 331

RESULT 10
US-09-902-540-3929
; Sequence 3929, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 3929
; LENGTH: 2943
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-3929

Query Match 6.3%; Score 40.6; DB 4; Length 2943;
Best Local Similarity 48.9%; Pred. No. 0.21;
Matches 109; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

QY 85 GTGCCCTCGAGACGGAGCGCTGCAGGCAACCCCATGAAGCTGCGTGCATCTCTCTGC 144
DB 1831 GTGCTGGAAGAGTGGAGCGCGCTGCGGAGCTGATGCGCGCGAGGCCCGC 1890

QY 145 ATGAAGAGAGAGGAGTGGAGGCCACACAGGTGGTGAATGGTTCTACAGGCCCGAGGC 204
DB 1891 ATGGCTACGGCAGGAGCTGAAGGCCCGCGTGGAGGGCGCGCCAGGTGAAGCGC 1950

QY 205 GGTAAAGATTTCCTTATTACAGTATCGAATGCCACAGGAGTGGAGAGCCCTTT 264
DB 1951 GCCTATGACCCGCTCCTGGACGTTCGACCGCGCGCGCTGGAGCGCTTGGTG 2010

QY 265 CAGGGGCGCTCGAGTGAATGGCAAGCAAGGACCTGCAGGAGC 307
DB 2011 AAGCGGCCAGGAGCGCATGGGCATCGAGGCCGAGGAGG 2053

RESULT 11
US-09-902-540-1194
; Sequence 1194, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1194
; LENGTH: 21295
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-1194
```

| Query Match | 6.3% | Score 40.6; | DB 4; | Length 21295; |
|-----------------------|-----------------|---|-----------|---------------|
| Best Local Similarity | 48.9%; | Pred. No. 0.54; | | |
| Matches 109; | Conservative 0; | Mismatches 114; | Indels 0; | Gaps 0; |
| Qy | 85 | GTGCCCTCGGAGACGAGCGCGTGCAGGCGCAACCCCATGAAGCTGCGCTGCATCTCTCTGC | 144 | |
| Db | 2609 | GTGCTGGAAGAGGTGGAGGCGCGCTGGCCGAGCTGGCGCTGATGCCGCGGAGGCCCGC | 2668 | |
| Qy | 145 | ATGAAGAGAGGAGGTGGAGGCCACCAACGGTGTGTGAATGGTTCTTACAGGCCCGAGGGC | 204 | |
| Db | 2669 | ATGCGCTTACGCGCAGGAGCTGAAGCGCCGCGTGGAGCGCGCGCCACGAGGTGAAGCGC | 2728 | |
| Qy | 205 | GGTAAAGATTTCCTTATTTACGAGTATCGGAATGCCACAGGAGGTGGAGAGCCCTTTT | 264 | |
| Db | 2729 | GCCTATGACCCGCTCTCTGGAGTGGCAGCTTTCACCCGCGCGGTGGAGCGCTGTGT | 2788 | |
| Qy | 265 | CAGGGCGCCTTGCAATGGAAATGGCAGCAAGCACTGCAGACG | 307 | |
| Db | 2789 | AAGCGCGCCAGGAGCGCATGGGATCGAGGCCACGAGAGG | 2831 | |

```

RESULT 12
US-09-902-540-7877
; Sequence 7877, Application US/09902540
; Patent NO. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16925
; SEQ ID NO 7877
; LENGTH: 1533
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
; US-09-902-540-7877

```

| Query Match | 6.2% | Score 40.4; | DB 4; | Length 1533; |
|-----------------------|-----------------|--|-----------|--------------|
| Best Local Similarity | 47.9% | Pred. No. 0.17; | | |
| Matches 116; | Conservative 0; | Mismatches 126; | Indels 0; | Gaps 0; |
| Qy | 232 | CGGAAATGGCCACAGAGAGGTGGAGAGCCCTTTTCAGGGGCGCCTGCAGTGGAAATGGCAGC | 291 | |
| Db | 310 | CGGATATCCCGCAGCAGTTGAGTCTCTCGAGGGAGGGCTCGTCTCGCGGATGGGGAC | 369 | |
| Qy | 292 | AAGGACCTGCAGGACGTTGCCATCAGTGTGCTCAACGTCTCAACGACACTCTGGGCCTC | 351 | |
| Db | 370 | ACGGAGCTGCAGGACCTGTGTGGCCCGGTACAGGTGACGCGGCTGGTCCAGTGGGC | 429 | |
| Qy | 352 | TACACCTGCAATGTGTCCCGGAGTTTGTAGTTTGAGGCGCATCGGCCCTTTGTGAAGACG | 411 | |
| Db | 430 | TTGACCGCACCCCATGCGCTGGGCTCGTGGAGGGGCGCTGGCGAACGTGGGCATCACC | 489 | |
| Qy | 412 | ACGGCGGTGATGCCCCCTAAGAGTCAACCGAGAGGCTGGAGAGACTTCACCTCTGTGGTC | 471 | |
| Db | 490 | GTGCTGTGTCGCTGCTCGTCTTCACCGCGCGGCGCTTCTCTGTGTACGCGCGCATCGTG | 549 | |
| Qy | 472 | TC | 473 | |
| Db | 550 | TC | 551 | |

RESULT 13
US-09-902-540-789
; Sequence 789, Application US/09902540
; Patent No. 6833447

```

; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 789
; LENGTH: 5228
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
; US-09-902-540-789

```

| Query Match | 6.2% | Score 40.4; | DB 4; | Length 5228; |
|-----------------------|-----------------|-----------------|---|--------------|
| Best Local Similarity | 47.94; | Pred. No. 0.31; | | |
| Matches 116; | Conservative 0; | Mismatches 126; | Indels 0; | Gaps 0; |
| Qy | 232 | CGGAATGCCAC | CAGGAGGTGGAGAGCCCTTTTCAGGGGGCCCTGCAGTGGAAATGGCAGC | 291 |
| | | | | |
| Db | 3057 | CGGGATGCCCG | CAGCAGTTGAGTGCTCTCCGAGGGAGGGCTCGTCTTCGCGGATGGGGAC | 3116 |
| | | | | |
| Qy | 292 | AAGGACCTGAG | ACGTCGTCCATCACTGTGCTCAACGTCACTCTGAACGACTCTGGGCTC | 351 |
| | | | | |
| Db | 3117 | ACGGAGCTGCA | GGACCTGTGTGGCGCCGGTACCAAGGTGACGCCGGCTGGGTCCACGTGGGC | 3176 |
| | | | | |
| Qy | 352 | TACACCTGCAAT | GTGTCCCGGGAGTTTGAGTTTGAGCGCGCATCGGCCCTTTGTGAAAGACG | 411 |
| | | | | |
| Db | 3177 | TTCCAGCGCAC | CCCATGCGCTGGGCTCTGGAAGGGGCGCTGGCGAACGTGGGCATCACC | 3236 |
| | | | | |
| Qy | 412 | ACGGGCTGATC | CCCCCTAAGAGTTCACGGAGAGGCTGGAGAGGACTTCACCTCTGTGGTC | 471 |
| | | | | |
| Db | 3237 | GTGCTGTGCG | CGTGCTCGTCTTACCGCGCGCGGCGCTTCTCTGTGCAGCGCGGCATCGTG | 3296 |
| | | | | |
| Qy | 472 | TC | 473 | |
| | | | | |
| Db | 3297 | TC | 3298 | |

RESULT 14
 US-09-902-540-3806
 ; Sequence 3806, Application US/09902540
 ; Patent No. 6833447
 ; GENERAL INFORMATION:
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Wiegand, Roger C.
 ; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
 ; FILE REFERENCE: 38-10(15849)B
 ; CURRENT APPLICATION NUMBER: US/09/902,540
 ; CURRENT FILING DATE: 2001-07-10
 ; PRIOR APPLICATION NUMBER: 60/217,883
 ; PRIOR FILING DATE: 2000-07-10
 ; NUMBER OF SEQ ID NOS: 16825
 ; SEQ ID NO 3806
 ; LENGTH: 732
 ; TYPE: DNA
 ; ORGANISM: Myxococcus xanthus
 US-09-902-540-3806

| | Query_Match | 6.0% | Score 39.2; | DB 4; | Length 732; |
|----|-----------------------|--|-----------------|-----------|-------------|
| | Best Local Similarity | 49.5%; | Pred. No. 0.27; | | |
| | Matches 101; | Conservative 0; | Mismatches 103; | Indels 0; | Gaps 0; |
| Qy | 297 | CTGTGAGGACGTTCATCACTGTCTCAACGTCATCTGTGAACGACATCTGGCCCTCTACAC | 356 | | |
| Db | 480 | CTGTGAGGACATCTCGAGACACGTGAAGATCGCCGGGCTCAAGCGCTGGGGGTCTTCAA | 539 | | |

| | | |
|-----|---|-----|
| 357 | CTGCNATGTCCTCCGGGAGTTTGATTTGAGCGGCATCGGCCCTTTGTCTGAAGACGACGCG | 416 |
| Qy | | |
| 540 | GTAACGCGCGCGCGAGCCCATGCTAAGTGTCTACCGCGACGACGCGTCGCGCG | 599 |
| Db | | |
| 417 | GCTGATCCCCCTTAAGAGTCAACGAGGAGGCTGGAGAGACTTCACTCTGTGGTCTTCAGA | 476 |
| Qy | | |
| 600 | GGTCAGACGTCCGTACTCAGCGCGCTGGCGAAGGTGGCTTCTCCGTGGAGGGCTTACCG | 659 |
| Db | | |
| 477 | AATCATGATGTACATCCTCTTGGT | 500 |
| Qy | | |
| 660 | GGAGAAGGTGGAGCGCTGCTGGT | 683 |
| Db | | |

RESULT 15

```

US-09-774-528-120
; Sequence 120, Application US/09774528
; Patent No. 6743619
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.
; APPLICANT: Yang, Yongdong
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wang, Dunrui
; APPLICANT: Drmanac, Radojke T.
; TITLE OF INVENTION: No. 6743619el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 802
; CURRENT APPLICATION NUMBER: US/09/774,528
; CURRENT FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 441
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 120
; LENGTH: 6202
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (49)..(4890)
US-09-774-528-120

```

| Query Match | 6.0%; | Score 39.2; | DB 4; | Length 6202; |
|-----------------------|-----------------|--|-----------|--------------|
| Best Local Similarity | 47.2%; | Pred. No. 0.76; | | |
| Matches 119; | Conservative 0; | Mismatches 133; | Indels 0; | Gaps 0; |
| Qy | 93 | GGAGACGGAGCGCGTGCAGGSCAACCCCATGAAGCTCGCTGCAATCTCTGCAATGAAGAG | 152 | |
| | | | | |
| Db | 2418 | GCACAAGGAGCTCGGGGAGAGGCAACGCTGACATGCTGAAGCGCAAGCGGAGCTGGAGGA | 2477 | |
| Qy | 153 | AGAGGAGGTGGAGGCCACACACGCTGTGTGGAATGTTCTACAGGCCCGAGGCGGTAAAGA | 212 | |
| | | | | |
| Db | 2478 | GCGGGAGAAAGTCTTTGACCACTGAGCGAGAGGCGCTGCACGAGGACAGAGGACAAACGC | 2537 | |
| Qy | 213 | TTTTCTTTATTTACAGAGTATCGGAATGCGCAACAGAGAGGTGAGAGCGCCCTTTTCAGGGGCG | 272 | |
| | | | | |
| Db | 2538 | CCTCGCCATGGCGGAGAAACAGAGGCTGCGGGCGAGCTGAGCAGGGETCAATTTCTCTGCA | 2597 | |
| Qy | 273 | CCTGCAGTGGATGGCAGCAGGACCTCGAGGAGTGTCCATCACTGTGCTCAACGTAC | 332 | |
| | | | | |
| Db | 2598 | CCACCACTGAAGGGGAGTACGAGGAGCTGCACGCCCAACCAAGAGGCTGAACACCTC | 2657 | |
| Qy | 333 | TCTGAACGACTC | 344 | |
| | | | | |
| Db | 2658 | ACTGAACAAACG | 2669 | |

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2005, 10:54:05 ; Search time 580.168 Seconds
(without alignments)
6768.525 Million cell updates/sec

Title: US-09-977-579-4_COPY_376_1023

Perfect score: 648

Sequence: 1 atgcctgcctcattagatt.....cggtagcagtgagggaatag 648

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5615251 seqs, 3030001701 residues

Total number of hits satisfying chosen parameters: 11230502

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*

3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*

4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*

5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*

6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*

7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*

8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*

9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*

10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*

11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*

12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*

13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*

14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*

15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*

16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*

17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*

18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*

19: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*

20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*

21: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

22: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|---------------------|
| 1 | 648 | 100.0 | 1261 | 11 | US-09-977-579-4 |
| 2 | 536.6 | 82.8 | 2220 | 11 | US-09-977-579-3 |
| 3 | 536.6 | 82.8 | 2632 | 13 | US-10-029-191-22 |
| 4 | 536.6 | 82.8 | 3108 | 13 | US-10-029-191-1 |
| 5 | 534.6 | 82.5 | 645 | 13 | US-10-029-191-21 |
| 6 | 148.2 | 22.9 | 657 | 13 | US-10-029-191-23 |
| 7 | 143.8 | 22.2 | 1335 | 18 | US-10-723-860-2247 |
| 8 | 142.2 | 21.9 | 1414 | 18 | US-10-477-272-1 |
| 9 | 141 | 21.8 | 407 | 17 | US-10-276-774-718 |
| 10 | 138.6 | 21.4 | 1490 | 9 | US-09-917-800A-1654 |
| 11 | 81.4 | 12.6 | 807 | 17 | US-10-401-916-12 |
| 12 | | | | | Sequence 4, Appli |
| 13 | | | | | Sequence 3, Appli |
| 14 | | | | | Sequence 22, Appli |
| 15 | | | | | Sequence 1, Appli |
| 16 | | | | | Sequence 21, Appli |
| 17 | | | | | Sequence 23, Appli |
| 18 | | | | | Sequence 2247, Ap |
| 19 | | | | | Sequence 1, Appli |
| 20 | | | | | Sequence 718, App |
| 21 | | | | | Sequence 1654, Ap |
| 22 | | | | | Sequence 12, Appli |

| | | | | | | |
|----|------|------|------|----|----------------------|--------------------|
| 12 | 81.4 | 12.6 | 974 | 17 | US-10-401-916-13 | Sequence 13, Appli |
| 13 | 49 | 7.6 | 243 | 16 | US-10-029-386-16214 | Sequence 16214, A |
| 14 | 49 | 7.6 | 569 | 16 | US-10-029-386-2514 | Sequence 2514, Ap |
| 15 | 48.6 | 7.5 | 3583 | 18 | US-10-723-860-6471 | Sequence 6471, Ap |
| 16 | 45.4 | 7.0 | 876 | 18 | US-10-767-701-10747 | Sequence 10747, A |
| 17 | 45.2 | 7.0 | 954 | 18 | US-10-425-115-10098 | Sequence 10098, A |
| 18 | 44.8 | 6.9 | 921 | 17 | US-10-369-493-43050 | Sequence 43050, A |
| 19 | 43.8 | 6.8 | 1434 | 17 | US-10-259-194A-622 | Sequence 622, App |
| 20 | 43.8 | 6.8 | 1720 | 17 | US-10-425-114-19213 | Sequence 19213, A |
| 21 | 43.8 | 6.8 | 1764 | 17 | US-10-425-114-15538 | Sequence 15538, A |
| 22 | 43.8 | 6.8 | 1794 | 17 | US-10-425-114-23057 | Sequence 23057, A |
| 23 | 43.8 | 6.8 | 2029 | 18 | US-10-425-115-15884 | Sequence 15884, A |
| 24 | 42.8 | 6.6 | 1298 | 18 | US-10-437-963-21288 | Sequence 21288, A |
| 25 | 41.6 | 6.4 | 1362 | 17 | US-10-282-122A-23492 | Sequence 23492, A |
| 26 | 40.2 | 6.2 | 779 | 17 | US-10-425-114-26163 | Sequence 26163, A |
| 27 | 39.8 | 6.1 | 1884 | 18 | US-10-437-963-53366 | Sequence 53366, A |
| 28 | 39.2 | 6.0 | 5973 | 17 | US-10-094-466-63 | Sequence 63, Appl |
| 29 | 39.2 | 6.0 | 6202 | 17 | US-10-120-988-120 | Sequence 120, Appl |
| 30 | 39 | 6.0 | 1798 | 18 | US-10-437-963-50269 | Sequence 50269, A |
| 31 | 38.6 | 6.0 | 922 | 18 | US-10-437-963-82867 | Sequence 82867, A |
| 32 | 38.6 | 6.0 | 1052 | 18 | US-10-425-115-154932 | Sequence 154932, A |
| 33 | 38.6 | 6.0 | 1179 | 9 | US-09-815-242-7698 | Sequence 7698, Ap |
| 34 | 38.6 | 6.0 | 1179 | 17 | US-10-282-122A-30011 | Sequence 30011, A |
| 35 | 38.6 | 6.0 | 1311 | 17 | US-10-369-493-42998 | Sequence 42998, A |
| 36 | 38.6 | 6.0 | 1536 | 17 | US-10-425-114-35508 | Sequence 35508, A |
| 37 | 38.6 | 6.0 | 2139 | 17 | US-10-369-493-32177 | Sequence 32177, A |
| 38 | 38.6 | 6.0 | 2720 | 18 | US-10-425-115-99348 | Sequence 99348, A |
| 39 | 38.4 | 5.9 | 3252 | 18 | US-10-437-963-102169 | Sequence 102169, A |
| 40 | 38.4 | 5.9 | 4143 | 18 | US-10-437-963-102168 | Sequence 102168, A |
| 41 | 38.2 | 5.9 | 1455 | 18 | US-10-767-701-14341 | Sequence 14341, A |
| 42 | 38 | 5.9 | 1689 | 16 | US-10-102-622-9 | Sequence 9, Appli |
| 43 | 38 | 5.9 | 1689 | 16 | US-10-102-622-11 | Sequence 11, Appli |
| 44 | 38 | 5.9 | 2692 | 17 | US-10-104-047-1554 | Sequence 1554, Ap |
| 45 | 37.6 | 5.8 | 1776 | 18 | US-10-437-963-42691 | Sequence 42691, A |

ALIGNMENTS

RESULT 1

US-09-977-579-4
; Sequence 4, Application US/09977579
; Publication No. US20040248240A1
; GENERAL INFORMATION:
; APPLICANT: Cambridge University Technical Services
; TITLE OF INVENTION: A novel family of beta sub-unit proteins from a voltage gated sod
; TITLE OF INVENTION: channel
; TITLE OF INVENTION: channel
; FILE OF INVENTION: nucleic acids encoding them and therapeutic or diagnostic uses th
; FILE REFERENCE: 674558-2001
; CURRENT APPLICATION NUMBER: US/09/977,579
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: PCT/EP00/01783
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60,129,473
; PRIOR FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1261
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-977-579-4

| | | | | | | | |
|-----------------------|--------|---|---------|------------|-----|--------|-------|
| Query Match | 100.0% | Score | 648; | DB | 11; | Length | 1261; |
| Best Local Similarity | 100.0% | Pred. No. | 1e-200; | | | | |
| Matches | 648; | Conservative | 0; | Mismatches | 0; | Indels | 0; |
| Gaps | 0; | | | | | | |
| QY | 1 | ATGCCTGCCTTCAATAGATTGTTTCCCTGGCTTCTCTCGTGTCTATCTACTGGTGCAGT | 60 | | | | |
| Db | 376 | ATGCCTGCCTTCAATAGATTGTTTCCCTGGCTTCTCTCGTGTCTATCTACTGGTGCAGT | 435 | | | | |
| QY | 61 | GTCTGTTCCCTGTGTGTGTGGAAGTGCCTTCGGACGGAGCCGTGACGGGCAACCCC | 120 | | | | |

Db 436 GTCTGCTTCCCTGTGTGTGGAAAGTGCCTCGGAGACGAGGCCGTGACGGCAACCC 495
QY 121 ATGAAGCTGCGTGCATCTCTGTCATGAAGAGAGAGAGGTGGAGCCCAACCGTGGTG 180
Db 496 ATGAAGCTGCGTGCATCTCTGTCATGAAGAGAGAGAGGTGGAGCCCAACCGTGGTG 555
QY 181 GAATGGTCTTACAGGCCCGGAGGCGGTAAAGATTTCTTATTACAGTATCGGAATGCG 240
Db 556 GAATGGTCTTACAGGCCCGGAGGCGGTAAAGATTTCTTATTACAGTATCGGAATGCG 615
QY 241 CACGAGAGGTGAGAGCCCTTTTCAGGGCGGCTGCAGTGGAAATGGCAGCAGACCTG 300
Db 616 CACGAGAGGTGAGAGCCCTTTTCAGGGCGGCTGCAGTGGAAATGGCAGCAGACCTG 675
QY 301 CAGGAGGTGTCATCTGTGCTCAACGTCACTCTGAACGACTCTGGCCTCTACACCTGC 360
Db 676 CAGGAGGTGTCATCTGTGCTCAACGTCACTCTGAACGACTCTGGCCTCTACACCTGC 735
QY 361 AATGTGTCGGGAGTTTGAATTTAGGGCGCATCGGCCCTTTGTGAAGACGACGGGCTG 420
Db 736 AATGTGTCGGGAGTTTGAATTTAGGGCGCATCGGCCCTTTGTGAAGACGACGGGCTG 795
QY 421 ATCCCTTAAGAGTCAACGAGGAGGTGAGAGGACTTCACTCTGTGCTCTCAGAAATC 480
Db 796 ATCCCTTAAGAGTCAACGAGGAGGTGAGAGGACTTCACTCTGTGCTCTCAGAAATC 855
QY 481 ATGATGTACATCTTCTGTCTTCTCAACCTGTGCTCTCATCGAGATGATATATTCG 540
Db 856 ATGATGTACATCTTCTGTCTTCTCAACCTGTGCTCTCATCGAGATGATATATTCG 915
QY 541 TACAGAAAGGTCTCAAAAGCGGAGGAGGAGCCCAAGAAACGCTGCTACCTTGGC 600
Db 916 TACAGAAAGGTCTCAAAAGCGGAGGAGGAGCCCAAGAAACGCTGCTACCTTGGC 975
QY 601 ATCCATCTGAGAACAGGAGAACTCTGGGTACCACTGAGTGGAGAAATAG 648
Db 976 ATCCATCTGAGAACAGGAGAACTCTGGGTACCACTGAGTGGAGAAATAG 1023

RESULT 2
US-09-977-579-3
; Sequence 3, Application US/0977579
; Publication No. US20040248240A1
; GENERAL INFORMATION:
; APPLICANT: Cambridge University Technical Services
; TITLE OF INVENTION: A novel family of beta sub-unit proteins from a voltage gated sodium channel
; TITLE OF INVENTION: channel
; FILE REFERENCE: nucleic acids encoding them and therapeutic or diagnostic uses thereof
; CURRENT APPLICATION NUMBER: US/09/977,579
; PRIOR FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: PCT/EP00/01783
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60,129,473
; PRIOR FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 2220
; TYPE: DNA
; ORGANISM: rat
US-09-977-579-3

Query Match 82.8%; Score 536.6; DB 11; Length 2220;
Best Local Similarity 89.3%; Pred. No. 3.1e-164;
Matches 578; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 1 ATGCTGCTTCAATAGATTGTTCCCTGGCTTCTCTGCTGCTTATCTACTGGGTGCT 60
Db 363 ATGCTGCTTCAACAGATTGTTCCCTAGCTTCTCTAGTGTCTATCTACTGGGTGCT 422
QY 61 GTCTGCTTCCCTGTGTGTGGAAAGTGCCTCGGAGACGAGGCCGTGACGGCAACCC 120

Db 423 GTCTGCTTCCCTGTGTGTGGAAAGTGCCTCGGAGACGAGGCCGTGACGGCAATCCC 482
QY 121 ATGAAGCTGCGTGCATCTCTGTCATGAAGAGAGAGAGGTGGAGCCCAACCGTGGTG 180
Db 483 ATGAAGCTGCGTGCATCTCTGTCATGAAGAGAGAGAGGTGGAGCCCAACCGTGGTG 542
QY 181 GAATGGTCTTACAGGCCCGGAGGCGGTAAAGATTTCTTATTACAGTATCGGAATGCG 240
Db 543 GAATGGTCTTACAGGCCCGGAGGCGGTAAAGATTTCTTATTACAGTATCGGAATGCG 602
QY 241 CACGAGAGGTGAGAGCCCTTTTCAGGGCGGCTGCAGTGGAAATGGCAGCAGACCTG 300
Db 603 CACGAGAGGTGAGAGCCCTTTTCAGGGCGGCTGCAGTGGAAATGGCAGCAGACCTG 662
QY 301 CAGGAGGTGTCATCTGTGCTCAACGTCACTCTGAACGACTCTGGCCTCTACACCTGC 360
Db 663 CAGGAGGTGTCATCTGTGCTCAACGTCACTCTGAACGACTCTGGCCTCTACACCTGC 722
QY 361 AATGTGTCGGGAGTTTGAATTTAGGGCGCATCGGCCCTTTGTGAAGACGACGGGCTG 420
Db 723 AATGTGTCGGGAGTTTGAATTTAGGGCGCATCGGCCCTTTGTGAAGACGACGGGCTG 782
QY 421 ATCCCTTAAGAGTCAACGAGGAGGTGAGAGGACTTCACTCTGTGCTCTCAGAAATC 480
Db 783 ATACCTTTGCGAGTCACTGAAGAGGCGGAGAGACTTCACTCTGTGCTCTCAGAAATC 842
QY 481 ATGATGTACATCTTCTGTCTTCTCAACCTGTGCTCTCATCGAGATGATATATTCG 540
Db 843 ATGATGTACATCTTCTGTCTTCTCAACCTGTGCTCTCATCGAGATGATATATTCG 902
QY 541 TACAGAAAGGTCTCAAAAGCGGAGGAGGAGCCCAAGAAACGCTGCTACCTTGGC 600
Db 903 TACAGAAAGGTCTCAAAAGCGGAGGAGGAGCCCAAGAAATGCGTCTGCTACCTTGGC 962
QY 601 ATCCATCTGAGAACAGGAGAACTCTGGGTACCACTGAGTGGAGAAATAG 647
Db 963 ATCCATCTGAGAACAGGAGAACTCTGGGTACCACTGAGTGGAGAAATAG 1009

RESULT 3
US-10-029-191-22
; Sequence 22, Application US/10029191
; Publication No. US20020160453A1
; GENERAL INFORMATION:
; APPLICANT: CURTIS, RORY A.J.
; TITLE OF INVENTION: NOVEL GENE ENCODING A SODIUM CHANNEL BETA-3 SUBUNIT
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 210147.00XX/501
; CURRENT APPLICATION NUMBER: US/10/029,191
; CURRENT FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 09/569,978
; PRIOR FILING DATE: 2000-05-12
; PRIOR APPLICATION NUMBER: US 60/134,198
; PRIOR FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 2632
; TYPE: DNA
; ORGANISM: Rattus sp.
US-10-029-191-22

Query Match 82.8%; Score 536.6; DB 13; Length 2632;
Best Local Similarity 89.3%; Pred. No. 3.3e-164;
Matches 578; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 1 ATGCTGCTTCAATAGATTGTTCCCTGGCTTCTCTGCTGCTTATCTACTGGGTGCT 60
Db 78 ATGCTGCTTCAACAGATTGTTCCCTAGCTTCTCTAGTGTCTATCTACTGGGTGCT 137
QY 61 GTCTGCTTCCCTGTGTGTGGAAAGTGCCTCGGAGACGAGGCCGTGACGGCAACCC 120
Db 138 GTCTGCTTCCCTGTGTGTGGAAAGTGCCTCGGAGACGAGGCCGTGACGGCAATCCC 197

| | | | |
|----|-----|---|-----|
| Qy | 121 | ATGAAGCTGGCGTCGATCTCCTGCATGAAGAGAGGAGGTGGAGCCACCACGGTGGT | 180 |
| Db | 198 | ATGAAGCTGGGTGCATCTCCTGCATGAAGAGGAGGAGGTGGAGGCCACCACCTGTGGT | 257 |
| Qy | 181 | GAATGGTTCTACAGGCCCGAGGGCGGTAAAGATTTCCTTATTTACAGTATCGGAATGGC | 240 |
| Db | 258 | GAGTGGTTCTACAGGCCTGAGGGCGGTAAAGATTTCCTATATATAGTATCGGAATGGC | 317 |
| Qy | 241 | CACAGAGGTGGAGAGCCCTTTCAGGGGGCGCTGCAGTGGAAATGGCAGCAGACCTG | 300 |
| Db | 318 | CACAGGAAGTGGAGAGCCCTTCCAAAGCCGCTGCGAGTGGAAATGGGAGCAAGACCTG | 377 |
| Qy | 301 | CAGGACGTGTCATCACTGTGCTCAAACGTCACTCTGAACGACTCTGGGCGCTTACACCTGC | 360 |
| Db | 378 | CAGGACGTATCCATCACTGTACTCAATGTCTACTTGAATGACTCTGGCGCTTACACATGC | 437 |
| Qy | 361 | AATGTGTCGGGGATTTGAGTTTGAGGCGGATCGGCCCTTTGTGAAGACGACGGCGTG | 420 |
| Db | 438 | AATGTGTCAGGGAGTTGCAATTGCGGCGACACAGGCCCTTTTGTGAAGACCAACGAGACTG | 497 |
| Qy | 421 | ATCCCCCTAAGAGTCACCGAGGAGCGTGGAGAGGACTTCACCTCTGTGGTCTCAGAAATC | 480 |
| Db | 498 | ATACCTTTGCGAGTCACTGAAGAGCGGGAGAGACTTCACCTCGTGGTCTCGGAATC | 557 |
| Qy | 481 | ATGATGTACATCTTCTGGTCTTCCCTACCCCTGTGGCTGCTCATCGAGATGATATTGCG | 540 |
| Db | 558 | ATGATGTACATCTCTCTGGTCTTCCCTCACCTTGTGGCTGTTTATTGAGATGATATTGCG | 617 |
| Qy | 541 | TACAGAAGGTCTCAAAGCCGAAGAGGCGACGCCAAGAAAAACGGCTCTGACTACCTTGGC | 600 |
| Db | 618 | TACAGAAGGTCTCTAAGGCCGAAGAGGCGACGACAGGAAATTCGGTCTGACTACCTTGGT | 677 |
| Qy | 601 | ATCCCATCTGAGAACAGAGGAACTCTCGGTACCAAGTGGAGGAATA | 647 |
| Db | 678 | ATCCCTTCAGAAACAAGGAACTCTGTGTAATCTGTGGAGGAATA | 724 |

RESULT 5
 US-10-029-191-21
 ; Sequence 21, Application US/10029191
 ; Publication No. US20020160453A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CURTIS, RORY A.J.
 ; TITLE OF INVENTION: NOVEL GENE ENCODING A SODIUM CHANNEL BETA-3 SUBUNIT
 ; TITLE OF INVENTION: PROTEIN
 ; FILE REFERENCE: 210147.00XX/5U1
 ; CURRENT APPLICATION NUMBER: US/10/029,191
 ; CURRENT FILING DATE: 2001-12-20
 ; PRIOR APPLICATION NUMBER: 09/569,978
 ; PRIOR FILING DATE: 2000-05-12
 ; PRIOR APPLICATION NUMBER: US 60/134,198
 ; PRIOR FILING DATE: 1999-05-14
 ; NUMBER OF SEQ ID NOS: 23
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 21
 ; LENGTH: 645
 ; TYPE: DNA
 ; ORGANISM: Rattus sp.
 US-10-029-191-21

| | Query Match | 82.5% | Score 534.6 | DB 13 | Length 645 |
|----|-----------------------|--|--------------------|----------|------------|
| | Best Local Similarity | 89.3% | Pred. No. 8.8e-164 | | |
| | Matches 576 | Conservative 0 | Mismatches 69 | Indels 0 | Gaps 0 |
| Qy | 1 | ATGCCTGCCCTCAATAGATTGTTTCCCTCGCTGCTTCTCTCGTGCCTATCTACTCTGGGTCACT | 60 | | |
| Db | 1 | ATGCCTGCCCTCAACAGATTGCTTCCCTCGCTGCTTCTCTAGTGCCTCATCTACTGGGTGAGA | 60 | | |
| Qy | 61 | GTCTGCTTCCCTGTGTGTGGAAAGTGCCCTCGAGACGGAGGCCGTGCAGGGCAACCC | 120 | | |
| Db | 61 | GTCTGCTTCCCTGTGTGTGGAAAGTGCCCTCGAGACAGAAAGCGGTGCAGGGCAATCCC | 120 | | |
| Qy | 121 | ATGAAGCTGGCGCTGCATCTCTCTGCATGAAGAGAGAGGAGGTGGAGGCCACCACCGTGGTG | 180 | | |

Db 121 ATGAAGCTGAGGTGCATCTCTGCTAAGAGGAGAGGTGGAGCCACCACTGTGGTG 180
QY 181 GAATGGTTCTACAGGCCGAGGCGGTAAAGATTTCTTTATTTACAGATATCGGAATGCG 240
Db 181 GAGTGGTTCTACAGGCCCTGAGGCGGTAAAGATTTCTTTATATATAGATATCGGAATGCG 240
QY 241 CACAGGAGGTGAGAGCCCTTTACAGGCGCGCTGCACTGGAATGAGAGGAGGACCTG 300
Db 241 CACAGGAGGTGAGAGCCCTTTCAAGGCGCTGCACTGGAATGAGAGGAGGACCTG 300
QY 301 CAGGAGGTGCATCACTGTCTCAAGCTCACTCTGAAAGAGCTCTGCGCTCTACACCTGC 360
Db 301 CAGGAGGTGCATCACTGTCTCAAGCTCACTCTGAAAGAGCTCTGCGCTCTACACCTGC 360
QY 361 AATGTGTCGGGAGTTGAGTTGAGGCGCATCGGCCCTTTGTGAAGACGAGCGGCTG 420
Db 361 AATGTGTCGGGAGTTGGAATTCGAGGCGACACAGGCGCTTTGTGAAGACGAGACTG 420
QY 421 ATCCCGCTAAGAGTACCGAGGAGGTGAGAGGACTTCACCTCTGTGTCTCAGAAATC 480
Db 421 ATACCTTTGCGAGTCACTGAAGAGGCGGAGAGACTTCACCTCGGTGTCTCGAAATC 480
QY 481 ATGATGTACATCTCTGTCTCTTCTCACCTGTGCTGCTCATCGAGATGATATTCG 540
Db 481 ATGATGTACATCTCTGTCTCTTCTCACCTGTGCTGCTTTATTGAGATGATATTCG 540
QY 541 TACAGAAAGGTCTCAAAAGCCGAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
Db 541 TACAGAAAGGTCTCTAAGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
QY 601 ATCCCATCTGAGACAGAGGAGACTCTGCGGTACCACTGAGGAGGAGGAGGAGGAG 645
Db 601 ATCCCTTACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 645

RESULT 6

US-10-029-191-23
; Sequence 23, Application US/10029191
; Publication No. US20020160453A1
; GENERAL INFORMATION:
; APPLICANT: CURTIS, RORY A.J.
; TITLE OF INVENTION: NOVEL GENE ENCODING A SODIUM CHANNEL BETA-3 SUBUNIT
; FILE REFERENCE: 210147.00XX/5U1
; CURRENT APPLICATION NUMBER: US/10/029,191
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 09/569,978
; PRIOR FILING DATE: 2000-05-12
; PRIOR APPLICATION NUMBER: US 60/134,198
; PRIOR FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 657
; TYPE: DNA
; ORGANISM: Lepus Sp.
US-10-029-191-23

Query Match 22.9%; Score 148.2; DB 13; Length 657;
Best Local Similarity 57.7%; Pred. No. 1.4e-37;
Matches 333; Conservative 0; Mismatches 223; Indels 21; Gaps 3;
QY 76 TGTGTGGAAGTCCCTCGGAGACGAGGCGGTGAGGCGGAGGAGGAGGAGGAGGAGGAGG 135
Db 61 TGGTGGAGTGGACTCGGAGACCGAAGCGGTGTACGGGATGACCTTCAAAATTTCTGTGC 120
QY 136 ATCTCTGTATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 195
Db 121 ATCTCTGAAAGCGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180
QY 196 CCGAGGCGGTAAAGATT-----TCCCTTATTACAGTATCGGAATGCGACCACTAG 246

Db 181 CAGAAAGGCACTGAGGAGTTGCTCAAGATCTCGCTATGAGAACGAGGTGCTGCAGCTG 240
QY 247 GAGGTGAGAGCCCTTTACAGGCGCGCTGCAAGTGAATGGCAG-----CAAGGAC 297
Db 241 GAGGAAGACGAGCGCTTTGAGGCGCGGTGTGTGGAACGCGCAGCGCGGCGCACCAAGGAC 300
QY 298 CTGAGGAGCGTGCATCACTGTGCTCAAGCTCACTCTGAAAGCACTCTGGCCTCTACACC 357
Db 301 CTGAGGAGCGTGCATCTTCACTCAATGCTCACTCACTCACTCACTCGGCGGACCTACCA 360
QY 358 TCAATATGTGTCGGGAGTTGAGTTGAGGCGCATCGGCCCTTTGTGAAGACGAGCGCG 417
Db 361 TGCCATGTCTACCGCTGCTCTCTTCGAAACTAGGAGCAACACACGAGCGTCTGTCAG 420
QY 418 CTGATCCCGCTTAAGAGTCAAGAGGAGGCTGAGAGGACTTCACCTCTGTGTCTCAGAA 477
Db 421 AAGATCCACCTGGAGTGTGGAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
QY 478 ATCATGATGTACATCTCTGCTCTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 537
Db 481 ATCATGATGTACATCTCTGCTCTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 540
QY 538 TCTACAGAAAGGTCTCAAAAGCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 594
Db 541 TCTACAGAAAGGTCTCGGCGCGCCACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
QY 595 CTGTCATCCCATCTGAGAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 631
Db 601 TTGGCCATCACCTCAGAAAGCAAGAAATTCGACGG 637

RESULT 7

US-10-723-860-2247
; Sequence 2247, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnick, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2247
; LENGTH: 1335
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-723-860-2247

Query Match 22.2%; Score 143.8; DB 18; Length 1335;
Best Local Similarity 56.9%; Pred. No. 5.1e-36;
Matches 334; Conservative 0; Mismatches 232; Indels 21; Gaps 3;
QY 66 CTTCTCTGTGTGTGGAAGTGCCTCGGAGACGAGGCGGTGAGGAGGAGGAGGAGGAGGAGG 125
Db 75 CTGCGGGGCTCGGTGGAGTGGACTCGGAGACCGAGGCGGTGTATGTGGATGACCTTCAA 134
QY 126 GCTGGCTGTCTCTCTGATGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 185
Db 135 AATTCTTTGCTCTCTCTGCAAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 194
QY 186 GTTCTACAGGCGCGGAGGCGGTAAAGATT-----TCCCTTATTACAGTATCGGAA 236
Db 195 GACCTTCGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 254
QY 237 TGGCCACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 290
Db 255 GTTGCAGCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 314

QY 291 ---CAAGGACCTGAGGAGTGTCCATCAGTGTGCTCAAGTCACTCTGAACGACTCTGG 347
 DB 315 CACCAAGAGCTGAGGAGTGTCTATCTTCACTACCAATGTCACTTACCAACCACTCGG 374
 QY 348 CCTCTACACCTGCAATGTGTCGGGAGTTGAGTTTGAAGCGCATCGGCCCTTTGTGAA 407
 DB 375 CGACTACAGTGGCCAGCTACCGCTGCTCTTCTTCCGAAACTACGAGCAACACACG 434
 QY 408 GACGACGGGCTGATCCCTTAAGATCACGAGGAGGCTGGAGAGACTTCACTCTGT 467
 DB 435 CGTCGTCAAGAAGATCCACATTTGAGGTAGTGACAAAGCCCAACAGAGACATGGCATCCAT 494
 QY 468 GGTCTCAGAAATCATGATGTACATCTTCTCTCACCCTGTGGTGTCTCATCGA 527
 DB 495 CGTGTCTGAGATCATGATGTATGTCTTGTGTGTGTGACCATATGGCTCTGGGAGA 554
 QY 528 GATGATATATTGTCTACAGAAAGGTCTCAAAAGCCGAAAGAG---GCAGCCCAAGAAAACGC 584
 DB 555 GATGATTTACTGTCTACAAAGAGATCGCTGCGCCACCGAGACTGCTGCACAGGAGAATGC 614
 QY 585 GTCTGACTACTTGGCCATCCCATCTGAGAACAGGAGAACTCTGCGG 631
 DB 615 CTCGGAATACCTGGCCATCACCTCTGAAGAAAGAAAGAACTGCAACG 661

RESULT 8

US-10-477-272-1
 ; Sequence 1, Application US/10477272
 ; Publication No. US20040191791A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Biomedics Limited
 ; TITLE OF INVENTION: P12
 ; FILE REFERENCE: SCN1B (R85C)
 ; CURRENT APPLICATION NUMBER: US/10/477,272
 ; CURRENT FILING DATE: 2003-11-10
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 1
 ; LENGTH: 1414
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-477-272-1

Query Match 21.9%; Score 142.2; DB 18; Length 1414;
 Best Local Similarity 56.7%; Pred. No. 1.7e-35;
 Matches 333; Conservative 0; Mismatches 233; Indels 21; Gaps 3;

QY 66 CTTCCCTGTGTGTGGAAGTGCCTCGGAGACGGAGCGGTGAGGGCAACCCCATGAA 125
 DB 151 CTGCGGGGCTGCGTGAGGTGGACTCGGAGACCGAGGCGGTGTATGGATGACCTTCAA 210
 QY 126 GCTGCGCTGTGATCTCTGATGAAGAGAGAGAGGTGGAGGCCACACGGTGTGGAAATG 185
 DB 211 AATTCTTTGTGATCTCTCTGCAAGCGCCGAGCGAGACCAACGCTGAGACCTTTCACCGAGTG 270
 QY 186 GTTCTACAGCCCGAGGCGGTAAAGATT-----TCCTTATTTACGATATCGAA 236
 DB 271 GACCTTCGCGCAGAAAGGCGACTGAGGAGTTTGTCAAGATCTGCGCTATGAGAAATGAGGT 330
 QY 237 TGGCCACAGAGGTGGAGAGCCCTTTTACGGGCGCTGCGAGTGGAAATGGCAG----- 290
 DB 331 GTTGACGCTGGAGGAGATGATGCTCTTCGAGGCGCGGTGTGTGGAAATGGCAGCGGG 390
 QY 291 ---CAAGGACCTGAGAGAGTGTCCATCATCTGTCTCAAGTCACTCTGAACGACTCTGG 347
 DB 391 CACCAAGAGCTGAGGAGTGTCTATCTTCACTACCAATGTCTACCTTACCAACCACTCGG 450
 QY 348 CCTCTACACCTGCAATGTGTCGGGAGTTTGAAGTGTGAGCGCATCGGCCCTTTGTGAA 407
 DB 451 CGACTACAGTGGCCAGCTGTACCGCTGCTCTTCTTCCGAAACTACGAGCAACACACG 510
 QY 408 GACGACGGGCTGATCCCTTAAGATCACCGAGGAGGCTGGAGAGGACTTCACTCTGT 467

DB 511 CGTCGTCAAGAAGATCCACATTTGAGGTAGTGGCAAAAGCCCAAGAGACATGGCATCCAT 570
 QY 468 GGTCTCAGAAATCATGATGTACATCTTCTGTGTTCTTCTCACCCTGTGGTGTCTCATCGA 527
 DB 571 CGTGTCTGAGATCATGATGTATGTGCTCATTTGTGTGTGTGACCATATGGCTCTGGCAGA 630
 QY 528 GATGATATATTGTCTACAGAAAGGTCTCAAAAGCCGAAAGAG---GCAGCCCAAGAAAACGC 584
 DB 631 GATGATTTACTGTCTACAAAGAGATCGCTGCGCCACCGAGACTGCTGCACAGGAGAATGC 690
 QY 585 GTCTGACTACTTGGCCATCCCATCTGAGAAACAGGAGAACTCTGCGG 631
 DB 691 CTCGGAATACCTGGCCATCACCTCTGAAGAAAGAAAGAACTGCAACG 737

RESULT 9

US-10-276-774-718/c
 ; Sequence 718, Application US/10276774
 ; Publication No. US20040053245A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hyseq, Inc.
 ; TITLE OF INVENTION: No. US20040053245A1el Nucleic Acids and Polypeptides
 ; FILE REFERENCE: 21272-030
 ; CURRENT APPLICATION NUMBER: US/10/276,774
 ; CURRENT FILING DATE: 2002-11-18
 ; PRIOR APPLICATION NUMBER: 09/560,875
 ; PRIOR FILING DATE: 2000-04-27
 ; PRIOR APPLICATION NUMBER: 09/496,914
 ; PRIOR FILING DATE: 2000-02-03
 ; NUMBER OF SEQ ID NOS: 2700
 ; SOFTWARE: Custom
 ; SEQ ID NO 718
 ; LENGTH: 407
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-276-774-718

Query Match 21.8%; Score 141; DB 17; Length 407;
 Best Local Similarity 96.6%; Pred. No. 2.7e-35;
 Matches 144; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 445 GCTGGAGAGGACTTCACTCTGTGGTCTCAGAAATCATGATGATCATCTTCTGTCTTC 504
 DB 230 GCTGGAGAGGACTTCACTCTGTGGTCTCAGAAATCATGATGATCATCTTCTGTCTTC 171
 QY 505 CTCAACCTGTGGTGTCTCATCGAGATGATATTTGCTACAGAAAGGTCTCAAAAGCCGAA 564
 DB 170 CTCACCTTGTGGTGTCTCATCGAGATGATATATTGCTACAGAAAGGTCTCAAAAGCCGAA 111
 QY 565 GAGGACGCCCAAGAAACGGTCTGACTA 593
 DB 110 GAGGACGCCCAAGAAACGGTAAAGTCCA 82

RESULT 10

US-09-917-800A-1654
 ; Sequence 1654, Application US/09917800A
 ; Patent No. US20020119462A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mendrick, Donna
 ; APPLICANT: Porter, Mark
 ; APPLICANT: Johnson, Kory
 ; APPLICANT: Castle, Arthur
 ; APPLICANT: Elashoff, Michael
 ; APPLICANT: Gene Logic, Inc.
 ; TITLE OF INVENTION: Molecular Toxicology Modeling
 ; FILE REFERENCE: 44921-5038-US
 ; CURRENT APPLICATION NUMBER: US/09/917,800A
 ; CURRENT FILING DATE: 2001-07-31
 ; PRIOR APPLICATION NUMBER: US 60/222,040
 ; PRIOR FILING DATE: 2000-07-31

PRIOR APPLICATION NUMBER: US 60/222,880
PRIOR FILING DATE: 2000-11-02
PRIOR APPLICATION NUMBER: US 60/290,029
PRIOR FILING DATE: 2001-05-11
PRIOR APPLICATION NUMBER: US 60/290,645
PRIOR FILING DATE: 2001-05-15
PRIOR APPLICATION NUMBER: US 60/292,336
PRIOR FILING DATE: 2001-05-22
PRIOR APPLICATION NUMBER: US 60/295,798
PRIOR FILING DATE: 2001-06-06
PRIOR APPLICATION NUMBER: US 60/297,457
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/298,884
PRIOR FILING DATE: 2001-06-19
PRIOR APPLICATION NUMBER: US 60/303,459
PRIOR FILING DATE: 2001-07-09
NUMBER OF SEQ ID NOS: 1740
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1654
LENGTH: 1490
TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM_017288
US-09-917-800A-1654

Query Match 21.4%; Score 138.6; DB 9; Length 1490;
Best Local Similarity 56.7%; Pred. No. 2.6e-34;
Matches 327; Conservative 0; Mismatches 229; Indels 21; Gaps 3;

QY 76 TGTGTGGAAGTCCCTCGAGCGGAGCGGTGAGGGCCACCCCATGAAGTGGCGCTCG 135
Db 280 TGGGTGGAGGTGAATCTGAGACCGGAGGAGTATGGGATGACCTTCAAAATCCTGTGT 339

QY 136 ATCTCTGTCATGAAGAGAGAGGAGGTGAGAGGCCACCAACCGTGTGGAATGGTTCTACAGG 195
Db 340 ATCTCTGTAAAGGTGTAAGTGAAGACCAACCGCGAGACCTTCACGGAGTGGACCTTCGCG 399

QY 196 CCGAGGCGGGTAAAGATT-----TCCTTATTAGAGTATCGGAATGGCCACCGAG 246
Db 400 CAGAAGGCGACAGAGGAATTTGTCAAGATCCTACGCTATGAGAATGAGGTGTGCGACGTG 459

QY 247 GAGGTGAGAGGCCCTTTCAGGGCGCGTGCAGTGAATGGCAG-----CAAGGAC 297
Db 460 GAGGAAGATGAGCGCTTTGAGGGCGCGTGTGTGTGAACGCTATGTCGGGCGACCAAGGAC 519

QY 298 CTGACGAGGTGTCCATCACTGTGCTCAAGCTCACTCTGAACGACTCTGGCCCTCTACACC 357
Db 520 CTGACGAGCTGTCCATCTTCATCAACCAATGTCAACCTACACCACTCTGGCGACTACGAA 579

QY 358 TGAATGTGTCCGGAGTTTGAAGTTTGAAGCGCATCGGCCCTTTGTGAAGAGCGCGG 417
Db 580 TGTCAAGTCTACCGTCTCCTCTTTGTATTAATTACGAGCACAACACCGCGTGTCAAG 639

QY 418 CTGATCCCTTAAGAGTCCAGGAGGCTGGAGAGACTTCACTCTGTGTCTCAGAA 477
Db 640 AGATCCACTGGAGTGTGGCAAGGCCAACAGAGATGATGCATCCATCGTGTGAGAG 699

QY 478 ATCATGATGTACATCTCTTGTGCTTCTCCACCTGTGGCTGTCTATCGAGATGATATAT 537
Db 700 ATCATGATGTACGTCTCATTTGGTGTAAACCATATGGCTGTGGCGAGATGGTGTAC 759

QY 538 TGCTACAGAAAGTCTCAAAAGCC-----GAAGGCGAGCCCAAGAAACCGGTCTGACTAC 594
Db 760 TGCTACAAGAAGATTGCTGTGTCGCGAAGCTGTGTCACAGAGAAATGCTCTCGGAATAC 819

QY 595 CTTGGCCATCCATCTCAGAACCAAGGAGAACTCTGGCG 631
Db 820 CTGGCCATTACTTCCGAGAGCAAGAGAACTGTACAG 856

RESULT 11
US-10-401-916-12

Sequence 12, Application US/10401916
Publication No. US20040002439A1
GENERAL INFORMATION:
APPLICANT: Qin, Ning
APPLICANT: Codd, Ellen
APPLICANT: D'Andrea, Michael
TITLE OF INVENTION: DNAs encoding human betala sodium channel subunit
FILE REFERENCE: ORT-1221
CURRENT APPLICATION NUMBER: US/10/401,916
CURRENT FILING DATE: 2003-03-28
PRIOR APPLICATION NUMBER: US/09/875,456A
PRIOR FILING DATE: 2001-09-10
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 12
LENGTH: 807
TYPE: DNA
ORGANISM: Homo sapiens
US-10-401-916-12

Query Match 12.6%; Score 81.4; DB 17; Length 807;
Best Local Similarity 57.5%; Pred. No. 1e-15;
Matches 195; Conservative 0; Mismatches 126; Indels 18; Gaps 2;

QY 66 CTTCCCTGTGTGTGGAAGTCCCTCGGAGACGAGGCGCGTGCAGGCAACCCCATGAA 125
Db 51 CTGCGGGGCTCGGTGGAGTGGACTCGGAGACCGGAGCCGTGTATGGGATGACCTTCAA 110

QY 126 GCTGGCGTGCATCTCTCGCATGAAGAGAGAGGAGGTGGAGGCCACCAACCGTGTGGAATG 185
Db 111 AATTCTTTCATCTCTCGCAAGCGCGCAGCGAGACCAACCGTGTGAGAAATG 170

QY 186 GTTCTACAGGCCGCGGCGGTAAGATT-----TCCTTATTAGAGTATCGGAA 236
Db 171 GACCTTCCGCCAAGAGGCACTGAGGAGTTTGTCAAGATCTCTGCGCTATGAGAAATGAGGT 230

QY 237 TGGCCACAGGAGGTGAGAGGCCCTTTTCAGGGCGCGTGCAGTGAATGGCAG----- 290
Db 231 GTTGCAGCTGAGAGAGATGAGCGCTTCGAGGCGCGGTGTGGAATGGCAGCGCGG 290

QY 291 ---CAAGGACCTGCGAGGAGTGTCCATCACTGTGCTCAACGTCACCTCTGAAACGACTCTGG 347
Db 291 CACCAAGACCTGCGAGGATCTGTCTATCTTCATCAACCAATGTCACTCAACCACTCGGG 350

QY 348 CTTCTACACTGCAATGTGTCCGGAGTTTGAGTTGA 386
Db 351 CGACTACGAGTGCACGCTACCGCTGCTCTTCTTCGA 389

RESULT 12

US-10-401-916-13
Sequence 13, Application US/10401916
Publication No. US20040002439A1
GENERAL INFORMATION:
APPLICANT: Qin, Ning
APPLICANT: Codd, Ellen
APPLICANT: D'Andrea, Michael
TITLE OF INVENTION: DNAs encoding human betala sodium channel subunit
FILE REFERENCE: ORT-1221
CURRENT APPLICATION NUMBER: US/10/401,916
CURRENT FILING DATE: 2003-03-28
PRIOR APPLICATION NUMBER: US/09/875,456A
PRIOR FILING DATE: 2001-09-10
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 13
LENGTH: 974
TYPE: DNA
ORGANISM: Homo sapiens
US-10-401-916-13

Query Match 12.6%; Score 81.4; DB 17; Length 974;
Best Local Similarity 57.5%; Pred. No. 1.1e-15;


```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (294)..(315)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1696)..(1712)
; OTHER INFORMATION: n is a, c, g, or t
US-10-723-860-6471

```

```

Query Match      7.5%; Score 48.6; DB 18; Length 3583;
Best Local Similarity 63.0%; Pred. No. 9.1e-05;
Matches 75; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY      454 GACTTCACCTCTGTGGTCTCAGAAATCATGATGTACATCCTTCTGGTCTTCTCACCCTG 513
Db      2229 GACATGGCATCCATCGTGTCTGAGATCATGATGTATGTGTCATTGGTGTGACCATA 2288

QY      514 TGGCTGCTCATCGAGATGATATATTGCTACAGAAAGGTCTCAAAAGCCGAAAGGCGAGC 572
Db      2289 TGGCTGCTGCGCAGAGATGATTTACTGCTACAGAGAGATGCTGCCGCCACGGAGAGCTGC 2347

```

Search completed: April 7, 2005, 15:19:35
Job time : 583.168 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2005, 08:50:34 ; Search time 3865.9 Seconds
(without alignments)
6380.306 Million cell updates/sec

Title: US-09-977-579-4_COPY_376_1023
Perfect score: 648
Sequence: 1 atgcctgccttcaatagatt.....cggtagcagtgagggaatag 648

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|-----------|----------------------|
| 1 | 646.4 | 99.8 | 648 | 9 | AY419145 | AY419145 Homo sapi |
| 2 | 646.4 | 99.8 | 2555 | 3 | CR609664 | CR609664 full-length |
| 3 | 646.4 | 99.8 | 4052 | 3 | HSMB01563 | AL136589 Homo sapi |
| 4 | 646 | 99.7 | 1052 | 5 | EX420015 | EX420015 BX420015 |
| 5 | 624.4 | 96.4 | 975 | 5 | EX445002 | EX445002 BX445002 |
| 6 | 546.8 | 84.4 | 582 | 5 | BP200910 | BP200910 BP200910 |
| 7 | 536.6 | 82.8 | 648 | 9 | AY419147 | AY419147 Mus muscu |
| 8 | 536.6 | 82.8 | 1359 | 3 | AK076466 | AK076466 Mus muscu |
| 9 | 536.6 | 82.8 | 3549 | 3 | AK049747 | AK049747 Mus muscu |
| 10 | 525.6 | 81.1 | 4149 | 3 | AK049286 | AK049286 Mus muscu |
| 11 | 518 | 79.9 | 636 | 5 | BM933157 | BM933157 UI-M-BH3- |
| 12 | 510.8 | 78.8 | 672 | 2 | BB614118 | BB614118 BB614118 |
| 13 | 508.2 | 78.4 | 4105 | 3 | BC058083 | BC058083 Mus muscu |
| 14 | 468 | 72.2 | 950 | 4 | BG294174 | BG294174 602391245 |
| 15 | 465 | 71.8 | 584 | 9 | AY419146 | AY419146 Pan trogl |
| 16 | 461.4 | 71.2 | 584 | 5 | BP361278 | BP361278 BP361278 |
| 17 | 447.6 | 69.1 | 723 | 6 | CA749311 | CA749311 UI-M-FY0- |
| 18 | 420.8 | 64.9 | 582 | 5 | BP311484 | BP311484 BP311484 |
| 19 | 408.4 | 63.0 | 927 | 5 | BQ713131 | BQ713131 AGENCOURT |
| 20 | 385.2 | 59.4 | 825 | 5 | BQ745919 | BQ745919 UI-M-EXO- |
| 21 | 382.4 | 59.0 | 846 | 6 | CD355879 | CD355879 UI-M-FY0- |
| 22 | 370.4 | 57.2 | 1069 | 5 | BM928131 | BM928131 AGENCOURT |
| 23 | 357.8 | 55.2 | 742 | 6 | CB526211 | CB526211 UI-M-FY0- |
| 24 | 351 | 54.2 | 986 | 5 | BU118914 | BU118914 603142291 |

| | | | | | | |
|----|-------|------|-----|---|----------|--------------------|
| 25 | 348.4 | 53.8 | 845 | 6 | CA327438 | CA327438 UI-M-FY0- |
| 26 | 345.2 | 53.3 | 692 | 7 | CN219870 | CN219870 WLA010B08 |
| 27 | 343.6 | 53.0 | 700 | 4 | BI739617 | BI739617 603361873 |
| 28 | 336.2 | 51.9 | 652 | 2 | BB652801 | BB652801 BB652801 |
| 29 | 330.6 | 51.0 | 892 | 5 | BU368614 | BU368614 603790206 |
| 30 | 326 | 50.3 | 595 | 5 | BU368531 | BU368531 603788820 |
| 31 | 310 | 47.8 | 450 | 6 | CB787935 | CB787935 AMGNNUC:N |
| 32 | 303.2 | 46.8 | 823 | 7 | CO428886 | CO428886 UI-M-HXO- |
| 33 | 302.8 | 46.7 | 734 | 7 | CK367344 | CK367344 AGENCOURT |
| 34 | 301.6 | 46.5 | 712 | 6 | CB526257 | CB526257 UI-M-FY0- |
| 35 | 299.6 | 46.2 | 780 | 5 | BQ770528 | BQ770528 UI-M-FY0- |
| 36 | 294.6 | 45.5 | 582 | 5 | BP202832 | BP202832 BP202832 |
| 37 | 291.6 | 45.0 | 750 | 6 | CD349206 | CD349206 UI-M-FY0- |
| 38 | 286 | 44.1 | 895 | 5 | BX743834 | BX743834 BX743834 |
| 39 | 281.6 | 43.5 | 714 | 7 | CF531573 | CF531573 UI-M-FY0- |
| 40 | 273 | 42.1 | 669 | 5 | BM951151 | BM951151 UI-M-RHO- |
| 41 | 266.4 | 41.1 | 471 | 6 | CB732717 | CB732717 AMGNNUC:N |
| 42 | 258.8 | 39.9 | 705 | 6 | CD349958 | CD349958 UI-M-FY0- |
| 43 | 255.8 | 39.5 | 798 | 7 | CF290312 | CF290312 AGENCOURT |
| 44 | 248.8 | 38.4 | 544 | 7 | CN666608 | CN666608 A0842F07- |
| 45 | 248.8 | 38.4 | 650 | 7 | CN665870 | CN665870 A0832C11- |

ALIGNMENTS

RESULT 1
AY419145
LOCUS
DEFINITION Homo sapiens HCM6793 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY419145
VERSION AY419145.1 GI:39775105
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 648)
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 648)
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT These sequences were made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..648
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>648
/locus_tag="HCM6793"
ORIGIN
Query Match 99.8%; Score 646.4; DB 9; Length 648;
Best Local Similarity 99.8%; Pred. No. 2.3e-163;
Matches 647; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGCCTGCCTTCAATAGATTGTTCCCTGGCTTCTCTCGTCTTACTCTGCTGGTCACT 60
DB 1 ATGCCTGCCTTCAATAGATTGTTCCCTGGCTTCTCTCGTCTTACTCTGCTGGTCACT 60

[illegible]

```
CONSRM The German cDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Medigenomix (Martinried/Germany) within the cDNA
sequencing Consortium of the German Genome Project.
This clone (DKFZp761F182) is available at the RZPD Deutsches
Reisourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp761F182
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.
FEATURES
Source
1..4052
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFZp761F182"
/db_xref="taxon:9606"
/clone="DKFZp761F182"
/tissue_type="amygdala"
/clone_lib="761 (synonym: hamy2). Vector pSport1; host
DH10B; sites NotI + SalI"
/dev_stage="adult"
/notes="voltage-gated sodium channel beta-3 subunit"
1..4052
/gene="DKFZp761F182"
804..1451
/gene="DKFZp761F182"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAB66524.1"
/db_xref="GI:1327681"
/db_xref="GOA:Q9NY72"
/db_xref="UniProt/Swiss-Prot:Q9NY72"
/translators="MPANRLFFLASLIYVSVCPVCEVPSSEAVQGNPKMLR
CISMKREVEATVVEWYFVRPEGKDFLIYEYRNGHQVEVSPFQGRLOWNGSKLDQD
VSLTVLNVLTNDGLVTCNVSREFEFEHAPFPVKTTLPLRVTEAGEDFTSVSEI
MMYLLVFLTLMLLIEMICYRKVKRAEAAQENASDYLALPSENKENSAPVEE"
ORIGIN
Query Match 99.8%; Score 646.4; DB 3; Length 4052;
Best Local Similarity 99.8%; Pred. No. 3.6e-163;
Matches 647; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGCCTCCCTTCAATAGATTGTTCCCTCGGCTTCTCTCGTCTATCTACTGGTCAGT 60
DB 804 ATGCCTCCCTTCAATAGATTGTTCCCTCGGCTTCTCTCGTCTATCTACTGGTCAGT 863
QY 61 GTCTGCTTCCCTGTGTGTGGAAAGTGCCTTCGAGACGAGCGCGTGCAGGGCAACCCC 120
DB 864 GTCTGCTTCCCTGTGTGTGGAAAGTGCCTTCGAGACGAGCGCGTGCAGGGCAACCCC 923
QY 121 ATGAAGTCGCGTGCATCTCTCGATGAAGAGAGAGAGGTGGAGGCCACACGGTGGTG 180
DB 924 ATGAAGTCGCGTGCATCTCTCGATGAAGAGAGAGAGGTGGAGGCCACACGGTGGTG 993
QY 181 GAATGTTTACAGGCCCGAGGGCGGTAAGATTCTTATTTACGAGTATCGAATGGC 240
DB 984 GAATGTTTACAGGCCCGAGGGCGGTAAGATTCTTATTTACGAGTATCGAATGGC 1043
QY 241 CACGAGAGGTGGAGAGCCCTTTTCAGGGCGCGCTGCAGTGGAAATGGCAGAGGACCTG 300
DB 1044 CACGAGAGGTGGAGAGCCCTTTTCAGGGCGCGCTGCAGTGGAAATGGCAGAGGACCTG 1103
QY 301 CAGGACGTGTCCATCACTGTGCTCAACGTCACTCTGAACGACTCTGGCCTCTACACCTGC 360
DB 1104 CAGGACGTGTCCATCACTGTGCTCAACGTCACTCTGAACGACTCTGGCCTCTACACCTGC 1163
QY 361 AATGTGTCCCGGAGTTTGAATTTAGCGCGCATCGGCCCTTTGTGAAGACGACGCGGCTG 420
DB 1164 AATGTGTCCCGGAGTTTGAATTTAGCGCGCATCGGCCCTTTGTGAAGACGACGCGGCTG 1223
```

```
QY 421 ATCCCTTAAGAGCTCACCGAGGAGGCTGGAGAGGACTTACCTCTGTGTCTCAGAAATC 480
DB 1224 ATCCCTTAAGAGCTCACCGAGGAGGCTGGAGAGGACTTACCTCTGTGTCTCAGAAATC 1283
QY 481 ATGATGTACATCCTTCTGGTCTTCTTCCCTCACCTCTGGCTGTCTCATCGAGATGATATATTC 540
DB 1284 ATGATGTACATCCTTCTGGTCTTCTTCCCTCACCTCTGGCTGTCTCATCGAGATGATATATTC 1343
QY 541 TACAGAAAGGTCTCAAAAGCCGAGAGGAGCCCAAGAAACGGCTGTGACTACCTTGGC 600
DB 1344 TACAGAAAGGTCTCAAAAGCCGAGAGGAGCCCAAGAAACGGCTGTGACTACCTTGGC 1403
QY 601 ATCCCATCTGAGAACCAAGAGAGAACTCTCGGTACCAGTGGAGGAATAG 648
DB 1404 ATCCCATCTGAGAACCAAGAGGAACTCTCGGTACCAGTGGAGGAATAG 1451

RESULT 4
BX420015
LOCUS BX420015 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
DEFINITION BX420015 1062 bp mRNA linear EST 01-MAY-2004
ACCESSION CS0DF023YA09 5-PRIME, mRNA sequence.
VERSION BX420015
KEYWORDS BX420015.2 GI:46929710
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Li, W.-B., Gruber, C., Jessee, J. and Polaves, D.
JOURNAL Full-length cDNA libraries and normalization
COMMENT Unpublished (2001)
On May 13, 2003 this sequence version replaced gi:30646738.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 6147.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?8=CS0DF023AA05QP1&c=6147.r.
FEATURES
source
1..1062
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF023YA09"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."
ORIGIN
Query Match 99.7%; Score 646; DB 5; Length 1062;
Best Local Similarity 99.7%; Pred. No. 3.3e-163;
Matches 646; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGCCTGCTTCAATAGATTGTTTCCCTGGGCTTCTCTCGTCTATCTACTGGTCAGT 60
DB 234 ATGCCTGCTTCAATAGATTGTTTCCCTGGGCTTCTCTCGTCTATCTACTGGTCAGT 293
QY 61 GTCTGCTTCCCTGTGTGTGGAAAGTGCCTTCGAGACGAGCGCGTGCAGGGCAACCCC 120
```

| | | |
|-----|---|-----|
| 294 | GTCTGTTCCCTGTGTGTGTGAGTAGTGCCCTTCGAGAGCGAGGGCCGTGCGAGGGCAACCCC | 353 |
| 121 | ATGAAGCTGCGCTGCATCTCTCTGCATGAAGAGAGAGAGGTGGAGGCCACACACGGTGGTG | 180 |
| 354 | ATGAAGCTGCGCTGCATCTCTCTGCATGAAGAGAGAGAGGTGGAGGCCACACACGGTGGTG | 413 |
| 181 | GAATGGTTTCTACAGGCCGAGGGCGGTAAAGATTTCTTATTTACGAGTATTCGGAATGGC | 240 |
| 414 | GAATGGTTTCTACAGGCCGAGGGCGGTAAAGATTTCTTATTTACGAGTATTCGGAATGGC | 473 |
| 241 | CACGAGAGTGGAGAGCCCTTTTCAGGGGGCGCTGCAGTGGATTCGACGACGAGCACTG | 300 |
| 474 | CACGAGAGTGGAGAGCCCTTTTCAGGGGGCGCTGCAGTGGATTCGACGACGAGCACTG | 533 |
| 301 | CAGGACGTGTCCATCACTGTGCTCAACGCTCACTCTGAAACGACTCTGGCCCTCTACACCTGC | 360 |
| 534 | CAGGACGTGTCCATCACTGTGCTCAACGCTCACTCTGAAACGACTCTGGCCCTCTACACCTGC | 593 |
| 361 | AATGTGTCCCGGAGTTTGAGTTTGAGGCGCATCGGCCCTTTGTGGAAGACGACGCGCTG | 420 |
| 594 | AATGTGTCCCGGAGTTTGAGTTTGAGGCGCATCGGCCCTTTGTGGAAGACGACGCGCTG | 653 |
| 421 | ATCCCCCTAAGATCACCGAGGAGGCTGGAGAGGACTTCACTCTGTGGTGTCTAGAAATC | 480 |
| 654 | ATCCCCCTAAGAGTCAACGAGGAGGCTGGAGAGGACTTCACTCTGTGGTGTCTAGAAATC | 713 |
| 481 | ATGATGTACATCTTCTGTGCTTCTCCTCACCTGTGGCTGCTCATCGAGATGATATATTC | 540 |
| 714 | ATGATGTACATCTTCTGTGCTTCTCCTCACCTGTGGCTGCTCATCGAGATGATATATTC | 773 |
| 541 | TACAGAAAGGTCTCAAAAGCCGAGAGGCGAGCCCAAGAAAAACGGCTCTGACTCTTGGCC | 600 |
| 774 | TACAGAAAGGTCTCAAAAGCCGAGAGGCGAGCCCAAGAAAAACGGCTCTGACTCTTGGCC | 833 |
| 601 | ATCCCATCTCAGAAACAGAGGAACTCTCGGTACCACTGGAGGGAATAG | 648 |
| 834 | ATCSGATCTCAGAAACAGAGGAACTCTCGGTACCACTGGAGGGAATAG | 881 |

| | |
|------------|---|
| RESULT | 5 |
| LOCUS | BX445002 |
| DEFINITION | 975 bp mRNA linear EST 04-MAY-2004 CS0D7023YA09 5-PRIME, mRNA sequence. |
| ACCESSION | EX445002 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone |
| VERSION | EX445002 |
| KEYWORDS | EST. |
| SOURCE | EX445002.2 GI:47009181 |
| ORGANISM | Homo sapiens (human) |
| REFERENCE | Homo sapiens |
| AUTHORS | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |
| TITLE | Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. |
| JOURNAL | 1 (bases 1 to 975) |
| COMMENT | Li,W.B., Gruber,C., Jesses,J. and Polayes,D. Full-length cDNA libraries and normalization Unpublished (2001) On May 15, 2003 this sequence version replaced gi:30782286. |

Genoscope - Centre national de séquençage
2 rue Gaston Crémieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five primer
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 6147.r
For more information about this cluster, see
<http://www.genoscope.cns.fr/cdna?s=CS1AF006ZE05QPl&c=6147.r>.

Location/Qualifiers
1..975
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

FEATURES
source

```

/clone="CS0DF023YA09"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/notes="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

```

ORIGIN

| Query Match | 96.4%; | Score 624.4; | DB 5; | Length 975; |
|-----------------------|-----------------|--|-----------|-------------|
| Best Local Similarity | 98.8%; | Pred. No. 2.2e-157; | | |
| Matches 636; | Conservative 3; | Mismatches 4; | Indels 1; | Gaps 1; |
| QY | 1 | ATGCTGCGCTTCAATAGATTGTTTCCCTGGCTTCTCTCGTGCTTATCTACTGGGTCAGT | 60 | |
| Db | 235 | ATGCTGCGCTTCAATAGATTGTTTCCCTGGCTTCTCTCGTGCTTATCTACTGGGTCAGT | 294 | |
| QY | 61 | GTCGCTTCCCTGTGTGTGTGGAAGTCCCTCGGAGCGAGGCCGTGCAGGGCAACCC | 120 | |
| Db | 295 | GTCGCTTCCCTGTGTGTGTGGAAGTCCCTCGGAGCGAGGCCGTGCAGGGCAACCC | 354 | |
| QY | 121 | ATGAAGCTGCGCTGCATCTCTCTGATCAAGAGAGAGAGGTGAGGCACCAACGGTGGTG | 180 | |
| Db | 355 | ATGAAGCTGCGCTGCATCTCTCTGATCAAGAGAGAGAGGTGAGGCACCAACGGTGGTG | 414 | |
| QY | 181 | GAATGGTTCTACAGGCCCGAGGGCGGTAAAGATTTCCTATTATTACAGTATCGGAATGCG | 240 | |
| Db | 415 | GAATGGTTCTACAGGCCCGAGGGCGGTAAAGATTTCCTATTATTACAGTATCGGAATGCG | 474 | |
| QY | 241 | CACGAGGAGGTGGAGAGCCCTTTTCAGGGGGCGCTCGACGTGGAATCGCAGCAGGACCTG | 300 | |
| Db | 475 | CACGAGGAGGTGGAGAGCCCTTTTCAGGGGGCGCTCGACGTGGAATCGCAGCAGGACCTG | 534 | |
| QY | 301 | CAGGACGTGTCATCTGCTCAACGCTCACTCTGAAACGACTCTGGCCCTCTACACCTGC | 360 | |
| Db | 535 | CAGGACGTGTCATCTGCTCAACGCTCACTCTGAAACGACTCTGGCCCTCTACACCTGC | 594 | |
| QY | 361 | AATGTGTCGGGAGTTTGATTTGAGCGCGATCGGCCCTTTGTGAAGACGACGGCGCTG | 420 | |
| Db | 595 | AATGAGTCCCGGAGTTTGATTTGAGCGCGCATCGGCCCTTTGTGAAGACGACGGCGCTG | 654 | |
| QY | 421 | ATCCCCCTTAAGAGTCAACGAGGAGGTGAGAGGACTTACACCTCTGTGGTCTCAGAAATC | 480 | |
| Db | 655 | ATCCCCCTTAAGAGTCAACGAGGAGGTGAGAGGACTTACACCTCTGTGGTCTCAGAAATC | 714 | |
| QY | 481 | ATGATGATACATCTTCTGGTCTTCCCTCAACCTGTGGTGTCTCATCGAGATGATATTGCG | 540 | |
| Db | 715 | ATGATGATACATCTTCTGGTCTTCCCTCAACCTGTGGTGTCTCATCGAGATGATATTGCG | 774 | |
| QY | 541 | TACGAAAGGTCTCAAAAGCCGAAGAGCGAGCCCAAGAAAACCGCTCTGACTACCTTGGC | 600 | |
| Db | 775 | TACGAAAGGTCTCAAAAGCCGAAGAGCGAGCCCAAGAAAACCGCTCTGACTACCTTGGC | 834 | |
| QY | 601 | ATCCCATCTGAGAACAAAGGAACTCTGCGGTACCAAGTGGAGA | 644 | |
| Db | 835 | ATCCCATCTGAGAACAAAGGAG-ACTCTGCGGTACAGTGGAGA | 877 | |

RESIST. 6

RECEIVED
BP200910

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE
ORGANIS

2445

REFERENCE

AUTHORS Suzuki, Y., Yamashita, R., Shirota, M., Sakakibara, Y., Chiba, J., Mizushima-Sugano, J., Nakai, K. and Sugano, S.

TITLE Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions

JOURNAL Genome Res. 14 (9), 1711-1718 (2004)

COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yeuzuki@ims.u-tokyo.ac.jp.

FEATURES

| | |
|--------|--|
| source | 1. 582 |
| | /organism="Homo sapiens" |
| | /mol_type="mRNA" |
| | /db_xref="taxon:9606" |
| | /clone="AMR03894" |
| | /tissue_type="amygdala" |
| | /clone_lib="Sugano CDNA library, amygdala" |

ORIGIN

| | | | | | | | |
|-----------------------|--------|--------------|---------------|------------|--------|--------|----|
| Query Match | 84.4%; | Score | 546.8; | DB 5; | Length | 582; | |
| Best Local Similarity | 99.6%; | Pred. | No. 1.8e-136; | | | | |
| Matches | 548; | Conservative | 0; | Mismatches | 2; | Indels | 0; |
| Gaps | 0; | | | | | | |

QY 3 GCCTGCCCTTCAATAGATGTTTCCCTTGGCTTCTCTCGTCTTATCTACTGGGTCACTGT 62

DB 33 GCCTGCCCTTCAATAGATGTTTCCCTTGGCTTCTCTCGTCTTATCTACTGGGTCACTGT 92

QY 63 CTGCTTCCCTGTGTGTGGAAGTGCCTCGGAGACGGAGCCGTGAGGGCACCCCAT 122

DB 93 CTGCTTCCCTGTGTGTGGAAGTGCCTCGGAGACGGAGCCGTGAGGGCACCCCAT 152

QY 123 GAAGCTCGCTGCATCTCCTGCATGAAGAGAGAGAGGTGTGAGGCCACACCGTGTGTGGA 182

DB 153 GAAGCTCGCTGCATCTCCTGCATGAAGAGAGAGAGGTGTGAGGCCACACCGTGTGTGGA 212

QY 183 ATGCTTCTACAGGCCCGAGGCGGTAAAGATTTCTCTTATTACAGTATCGGAATGGCCA 242

DB 213 ATGCTTCTACAGGCCCGAGGCGGTAAAGATTTCTCTTATTACAGTATCGGAATGGCCA 272

QY 243 CCAGGAGTGGAGAGCCCTTTACGGGGCGCTGCAGTGAATGGCAGCAGACCTTGCA 302

DB 273 CCAGGAGTGGAGAGCCCTTTACGGGGCGCTGCAGTGAATGGCAGCAGACCTTGCA 332

QY 303 GGCAGTGTCCATCACTGTGTCTCAAGTCACTCTGAACGACTCTGGCTCTACACCTGCAA 362

DB 333 GGCAGTGTCCATCACTGTGTCTCAAGTCACTCTGAACGACTCTGGCTCTACACCTGCAA 392

QY 363 TGTGTCCCGGAGTTTGAAGTTCAGTTCGAGCGCATCGGCCCTTTGTGAAGACGACGGGCTGT 422

DB 393 TGTGTCCCGGAGTTTGAAGTTCAGTTCGAGCGCATCGGCCCTTTGTGAAGACGACGGGCTGT 452

QY 423 CCCCTTAAGAGTCAACGAGAGGCTGAGAGGACTTCACTCTGTGTCTCAGAAATCAT 482

DB 453 CCCCTTAAGAGTCAACGAGAGGCTGAGAGGACTTCACTCTGTGTCTCAGAAATCAT 512

QY 483 GATGTACATCCTTCTGTCTTCTTCAACCTGTGTGCTCTCATCGAGATGATATTGCTTA 542

DB 513 GATGTACATCCTTCTGTGTCTTCTTCACTGTGTGCTCTCATCGAGATGATATTGCTTA 572

QY 543 CAGAAAGGTC 552

DB 573 CAGAAAGGTC 582

RESULT 7

AV419147 648 bp DNA linear GSS 12-DEC-2003

LOCUS Mus musculus HCM6793 gene, VIRTUAL TRANSCRIPT, partial sequence,

DEFINITION genomic survey sequence.

ACCESSION AV419147

VERSION AV419147.1

KEYWORDS GI:39775107 GSS.

Db 541 TACAGAAAGTCTCTAAGCCGAGAGGAGCTCAGAAATCGCTGACTACCTTGT 600
|||||
Qy 601 ATCCCATCTGAGAACAGGAGAACTCTCGGTTACAGTGGAGGAATA 647
|||||
Db 601 ATCCCTTCAGAGAACAGGAGAACTCTGTGGTACCGTGGAGGAATA 647
|||||
RESULT 8
AK076466 1359 bp mRNA linear HTC 03-APR-2004
LOCUS Mus musculus 0 day neonate head cDNA, RIKEN full-length enriched
DEFINITION library, clone:4833414B02 product:VOLTAGE-GATED SODIUM CHANNEL
BETA-3 SUBUNIT (SODIUM CHANNEL BETA 3 SUBUNIT) homolog [Rattus
norvegicus], full insert sequence.
ACCESSION AK076466
VERSION AK076466.1 GI:26345409
KEYWORDS HTC; CAP trapper;
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kusunagi, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multichannel sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6
AUTHORS Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,
Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,
Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,
Suzuki, K., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,
Tejima, Y., Toba, T., Yamamura, T., Yamanaka, I., Yasunishi, A.,
Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

TITLE Direct Submission
JOURNAL Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
FEATURES
Location/Qualifiers
1..1359
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CS7BL/6J"
/db_xref="FANTOM:DB:4833414B02"
/db_xref="taxon:10090"
/clone="4833414B02"
/tissue_type="head"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="0 day neonate"
57..704
CDS
/note="unnamed protein product; VOLTAGE-GATED SODIUM
CHANNEL BETA-3 SUBUNIT (SODIUM CHANNEL BETA 3 SUBUNIT)
homolog [Rattus norvegicus] (SPTR|Q9JK00, evidence: FASTY,
99.5%ID, 100%length, match=646)
putative"
/codon_start=1
/protein_id="BAC36356.1"
/db_xref="GI:26345410"
/translation="MPAFNRLPLASLVLYVVRVCFVCPVSETEAVQNSMKLR
CISCMKREVAATTVENFYRPEGKDFLIYRNHGVESPFQGRLOWNSKDLQD
VSIIVNTLNDLSGLYTCNSREFEFAHRFPVKTRILPLRVTEAGEDFTSVVSEI
MMYILLVLTLLWLFIEIMTYCYRKRKSKAEAAQENASDYLAIPSENKENSVPVEE"
ORIGIN
Query Match 82.8%; Score 536.6; DB 3; Length 1359;
Best Local Similarity 89.3%; Pred. No. 1.2e-133; Indels 0; Gaps 0;
Matches 578; Conservative 0; Mismatches 69;
Qy 1 ATGCGTGCCTTCAATAGATTGTTTCCCTGGGTTCTCTGCTCTTATCTACTGGGTCACT 60
|||||
Db 57 ATGCGTGCCTTCAACAGATTGCTTCCCTAGCTTCTTAGTGCTCATCTACTGGGTCA 116
|||||
Qy 61 GTCTGCTTCCCTGTGTGTGGAAAGTGCCTCGGAGAGCGGCGGTGAGGCAACCCC 120
|||||
Db 117 GTCTGCTTCCCTGTGTGTAGAGTACCTCGGAGACAGAGCCGTGAGGCAATTC 176
|||||
Qy 121 ATGAGAGCTGCGTGCATCTCTCGCATGAGAGAGGAGTGGAGCCACCACTGGGTG 180
|||||
Db 177 ATGAGAGCTGAGATGATCTCTCGCATGAGAGGAGGAGTGGAGCCACCACTGATG 236
|||||
Qy 181 GAATGCTTCTACAGGCCCGAGGCGGTAAAGATTTCTTTATTTACGAGTATCGAATGC 240
|||||
Db 237 GAGTGTCTACAGGCTGAGGCGGTAAAGATTTCTTTATTTATGATATCGAATGC 296
|||||
Qy 241 CACAGAGGAGTGGAGAGCCCTTTCAGGGCCCTGTCAGTGAATGGAGGAGGACCTG 300
|||||
Db 297 CACAGAGGAGTGGAGAGCCCTTTCAGAGTCTGTCAGTGAATGGAGGAGGACCTG 356
|||||
Qy 301 CAGGAGCTGCTCATCTCACTGTGCTCAACGCTCTCTGAACGACTCTGGCCCTCTACACCTGC 360
|||||
Db 357 CAGGAGTATTCATCTCACTGTGCTCAATGATGATCTGAGTCTCTGGCCCTCTACACATCT 416
|||||
Qy 361 AATGTGTCCCGGAGTTTGAAGTGGAGCGCATCGGCCCTTTGTGAAGACGACGGGCTG 420
|||||
Db 417 AATGTGTCCCGGAGTTTGAAGTGGAGCGCATCGGCCCTTTGTGAAGACGACGACTA 476
|||||


```
Db 472 CACCAGAGGTGGAGGCCCTTCCAAAGTTCGTCTGCAATGGGAGAAAGACCTG 531
QY 301 CAGGAGCTGTCCATCATCTGTCTCAACGTCACCTCTGAAGACTCTGGCTCTACACCTGC 360
Db 532 CAGGAGTATCCATCACTGTCTCATGTGTCACCTCTGAATGACTCTGGCTCTACACATGT 591
QY 361 AATGTCTCCGGAGTTTGAGTTTGAGGCGCATCGGCCCTTTGTGTGAAGACGCGGGCTG 420
Db 592 AATGTCTCAGGAGTTTGAGTTTGAAGACACACCGGCCCTTTGTGTGAAGACCAAGACTA 651
QY 421 ATCCCCCTAAGAGTCAACGAGGAGGCTGAGAGGACTTCACCTCTGTGTCTCAGAAATC 480
Db 652 ATACCCCTCGGAGTCACTGAAGAGGCGGAGAGACTTCACCTCGGTGTCTCGGAAATC 711
QY 481 ATGATGTACATCTCTCTGTCTTCTCCACCTCTGTGCTGCTCATCGAGATGATATATTC 540
Db 712 ATGATGTACATCTCTCTGTCTTCTCTCTCACTGTGTGCTGTTTATTGAGATGATATTC 771
QY 541 TACAGAAAGTCTCAAAAGCCGAGAGGAGCGCCCAAGAAACGCGTCTGACTACCTTCCC 600
Db 772 TACAGAAAGTCTCTAAGCCGAGAGGAGCGCTCAGGAAATCGCTGACTACCTTGGT 831
QY 601 ATCCCATCTGAGNACAGAGGAGTCTGCGGTACCACTGAGGAGATA 647
Db 832 ATCCCTTCAGAGAACCAAGGAGAACTCTGTGTACCGCTGGAGGAATA 878

RESULT 10
AC049286
LOCUS
DEFINITION
Mus musculus ES cells cDNA, RIKEN full-length enriched library,
clone: C330019103 product: VOLTAGE-GATED SODIUM CHANNEL BETA-3
SUBUNIT (SODIUM CHANNEL BETA 3 SUBUNIT) homolog [Rattus
norvegicus], full insert sequence.
AC049286
VERSION
AK049286.1 GI:26093400
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
MEDLINE
10349636
PUBMED
REFERENCE
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
MEDLINE
11042159
PUBMED
REFERENCE
3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujisawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanishi, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
MEDLINE
11076861
PUBMED
REFERENCE
4 The RIKEN Genome Exploration Research Group Phase II Team and the
PANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
TITLE
```

```
JOURNAL
REFERENCE
5 Nature 409, 685-690 (2001)
AUTHORS
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL
REFERENCE
6 Nature 420, 563-573 (2002)
AUTHORS
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanaoka, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, P., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, T., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
TITLE
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
COMMENT
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
FEATURES
source
1..4149
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM DB:C330019103"
/db_xref="taxon:10090"
/clone="C330019103"
/cell_type="ES cells"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
misc_feature
410..1058
/note="VOLTAGE-GATED SODIUM CHANNEL BETA-3 SUBUNIT (SODIUM
CHANNEL BETA 3 SUBUNIT) homolog [Rattus norvegicus]
(SPTR|Q9JUK0, evidence: FASTY, 99.5%ID, 100%length,
match=646)
putative"
polya_signal
4129..4134
/note="putative"
polya_site
4149
/note="putative"
ORIGIN
Query Match 81.1%; Score 525.6; DB 3; Length 4149;
Best Local Similarity 89.2%; Pred. No. 1.5e-130;
Matches 578; Conservative 0; Mismatches 69; Indels 1; Gaps 1;
QY 1 ATGCCTGCTTCAATAGATGTTTCCCTGGCTTCTCTGCTTATCTACTGGGTCACT 60
Db 410 ATGCCTGCTTCAACAGATGTTTCCCTAGTCTTCTAGTGTCTATCTACTGGGTCA 469
QY 61 GTCCTGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 119
Db 470 GTCCTGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 529
QY 120 CATGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 179
Db 530 CATGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 589
```



```

Db      547  TACAGAAAGGCTCTTAAGTCGAGAGGCGAGCTCAGAAAATCGCTGACTACCTTGCT 606
QY      601  ATCCCATCTGAGAACAGGAGACTCTGCG 630
Db      607  ATCCCTTCAGAGAACACAGGAACTCTGTG 636

RESULT 12
BB614118 672 bp mRNA linear EST 26-OCT-2001
LOCUS    BB614118 RIKEN full-length enriched, 0 day neonate head Mus
DEFINITION musculus cDNA clone 4833414B02 5', mRNA sequence.
ACCESSION BB614118
VERSION    BB614118.1 GI:16454578
KEYWORDS  EST.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
REFERENCE 1 (bases 1 to 672)
AUTHORS   Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,
Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,
Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,
Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
Takeda,Y., Tanaka,T., Toyota,T., Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamana,K.I.,
Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
Hayashizaki,Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
e mouse tissues.
Location/Qualifiers
1. 672
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="4833414B02"
/sex="mixed"
/tissue_type="head"
/dev_stage="0 day neonate"
/lab_host="DH10B"

FEATURES
source
BC058083 4105 bp mRNA linear HTC 19-NOV-2003
LOCUS

ORIGIN
Query Match 78.8%; Score 510.8; DB 2; Length 672;
Best Local Similarity 89.2%; Pred. No. 9.6e-127;
Matches 551; Conservative 0; Mismatches 67; Indels 0; Gaps 0;
QY 1 ATGCGTCGCTTCATAGATTGTTCCCTGGCTTCTCTCGTGTCTTACTTGGGTCACT 60
Db 55 ATGCGTCGCTTCAACAGATTGTTCCCTGGCTTCTCTCGTGTCTTACTTGGGTCA 114
QY 61 GTCTGCTTCCCTGTGTGTGGAGTGCCTCGAGAGCGAGGCGCTGAGGGCAACCCC 120
Db 115 GTCTGCTTCCCTGTGTGTGGAGTGCCTCGAGAGCGAGGCGCTGAGGGCAATTC 174
QY 121 ATGAAGCTGCGCTGCATCTCTGTCATGAAGAGAGAGAGGTGGAGGCCACCACTG 180
Db 175 ATGAAGCTGAGATGCATCTCTGTCATGAAGAGAGAGGTGGAGGCCACCACTG 234
QY 181 GAATGTTCTACAGCCCGAGGGCGGTAAAGATTTCCTTATTACGAGTATCGGAATGC 240
Db 235 GAGTGGTTCACAGCCCTGAGGGCGGTAAAGATTTCCTTATATGATGATATCGAAATGC 294
QY 241 CACAGAGAGGTGGAGAGCCCTTTTCAGGGCGCTTCAGTGGAAATGGCAGAACCTG 300
Db 295 CACAGAGAGGTGGAGAGCCCTTCCAGGTGCTCTGCAATGGAAATGGAGCAAGCTG 354
QY 301 CAGGACGTGTCATCATCTGTGCTCAACCTCTCTGAAAGAGCTCTGGCTCTTACACCTGC 360
Db 355 CAGGACGTATCCATCACTGTCTCAATGTCACTCTGAAATGACTCTGGCTCTTACACATGT 414
QY 361 AATGTGTCGGGAGTTTGAGTTTGAGGGCGCATCGGCCCTTTGTGAAGACGAGCGGGCTG 420
Db 415 AATGTGTCAGGGAGTTTGAGTTTGAGGACACACCGGCCCTTTGTGAAGACCAAGACTA 474
QY 421 ATCCCTCTAAGAGTCAACGAGAGGCTGCGAGAGGACTTTCACCTCTGTGGTCTCGAAATC 480
Db 475 ATACCTCTCGAGTCACTGAAGAGGCGGAGAGACTTTCACCTCGTGGTCTCGGAATC 534
QY 481 ATGATGTAATCTCTCTGTGCTTCTTCTCACTCTGTGGTGTGCTCATCGAGATATATTGC 540
Db 535 ATGATGTAATCTCTCTGTGCTTCTTCTCACTCTGTGGTGTGCTTATTTAGATGATCTATTGC 594
QY 541 TACAGAAAGGCTTCAAAAGCCGAGAGGCGAGCCCAAGAAACGGCTCTGACTTACCTTGC 600
Db 595 TACAGAAAGGCTTCAAGCCGAGAGGCGAGCTCAAGAAAAGCCCTCTGACTTACCTTGT 654
QY 601 ATCCCATCTGAGAACAG 618
Db 655 ATCCCTTCAGAGAACAG 672

RESULT 13
BC058083
LOCUS

```

547 TACAGAAAGGCTCTTAAGTCGAGAGGCGAGCTCAGAAAATCGCTGACTACCTTGCT 606

601 ATCCCATCTGAGAACAGGAGACTCTGCG 630

607 ATCCCTTCAGAGAACACAGGAACTCTGTG 636

RESULT 12

BB614118 672 bp mRNA linear EST 26-OCT-2001
 LOCUS BB614118 RIKEN full-length enriched, 0 day neonate head Mus
 DEFINITION musculus cDNA clone 4833414B02 5', mRNA sequence.

ACCESSION BB614118

VERSION BB614118.1 GI:16454578

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 672)
 AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,
 Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,
 Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,
 Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,
 Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
 Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
 Takeda,Y., Tanaka,T., Toyota,T., Muramatsu,M. and Hayashizaki,Y.
 RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
 Unpublished (2001)

REFERENCE

AUTHORS

Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,
 Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,
 Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,
 Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,
 Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
 Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
 Takeda,Y., Tanaka,T., Toyota,T., Muramatsu,M. and Hayashizaki,Y.
 RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
 Unpublished (2001)
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center(GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/

TITLE

JOURNAL

COMMENT

Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
 Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
 Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
 Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
 and Hayashizaki,Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
 Sugahara,Y. and Hayashizaki,Y.
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamana,K.I.,
 Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
 Hayashizaki,Y.
 Computational Analysis of Full-Length Mouse cDNAs Compared with
 Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
 Please visit our web site (http://genome.gsc.riken.go.jp) for
 further details.

e mouse tissues.

Location/Qualifiers

1. 672

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="4833414B02"

/sex="mixed"

/tissue_type="head"

/dev_stage="0 day neonate"

/lab_host="DH10B"

FEATURES

source

BC058083 4105 bp mRNA linear HTC 19-NOV-2003
 LOCUS
 ORIGIN
 Query Match 78.8%; Score 510.8; DB 2; Length 672;
 Best Local Similarity 89.2%; Pred. No. 9.6e-127;
 Matches 551; Conservative 0; Mismatches 67; Indels 0; Gaps 0;
 QY 1 ATGCGTCGCTTCATAGATTGTTCCCTGGCTTCTCTCGTGTCTTACTTGGGTCACT 60
 Db 55 ATGCGTCGCTTCAACAGATTGTTCCCTGGCTTCTCTCGTGTCTTACTTGGGTCA 114
 QY 61 GTCTGCTTCCCTGTGTGTGGAGTGCCTCGAGAGCGAGGCGCTGAGGGCAACCCC 120
 Db 115 GTCTGCTTCCCTGTGTGTGGAGTGCCTCGAGAGCGAGGCGCTGAGGGCAATTC 174
 QY 121 ATGAAGCTGCGCTGCATCTCTGTCATGAAGAGAGAGGTGGAGGCCACCACTG 180
 Db 175 ATGAAGCTGAGATGCATCTCTGTCATGAAGAGAGAGGTGGAGGCCACCACTG 234
 QY 181 GAATGTTCTACAGCCCGAGGGCGGTAAAGATTTCCTTATTACGAGTATCGGAATGC 240
 Db 235 GAGTGGTTCACAGCCCTGAGGGCGGTAAAGATTTCCTTATATGATGATATCGAAATGC 294
 QY 241 CACAGAGAGGTGGAGAGCCCTTTTCAGGGCGCTTCAGTGGAAATGGCAGAACCTG 300
 Db 295 CACAGAGAGGTGGAGAGCCCTTCCAGGTGCTCTGCAATGGAAATGGAGCAAGCTG 354
 QY 301 CAGGACGTGTCATCATCTGTGCTCAACCTCTCTGAAAGAGCTCTGGCTCTTACACCTGC 360
 Db 355 CAGGACGTATCCATCACTGTCTCAATGTCACTCTGAAATGACTCTGGCTCTTACACATGT 414
 QY 361 AATGTGTCGGGAGTTTGAGTTTGAGGGCGCATCGGCCCTTTGTGAAGACGAGCGGGCTG 420
 Db 415 AATGTGTCAGGGAGTTTGAGTTTGAGGACACACCGGCCCTTTGTGAAGACCAAGACTA 474
 QY 421 ATCCCTCTAAGAGTCAACGAGAGGCTGCGAGAGGACTTTCACCTCTGTGGTCTCGAAATC 480
 Db 475 ATACCTCTCGAGTCACTGAAGAGGCGGAGAGACTTTCACCTCGTGGTCTCGGAATC 534
 QY 481 ATGATGTAATCTCTCTGTGCTTCTTCTCACTCTGTGGTGTGCTCATCGAGATATATTGC 540
 Db 535 ATGATGTAATCTCTCTGTGCTTCTTCTCACTCTGTGGTGTGCTTATTTAGATGATCTATTGC 594
 QY 541 TACAGAAAGGCTTCAAAAGCCGAGAGGCGAGCCCAAGAAACGGCTCTGACTTACCTTGC 600
 Db 595 TACAGAAAGGCTTCAAGCCGAGAGGCGAGCTCAAGAAAAGCCCTCTGACTTACCTTGT 654
 QY 601 ATCCCATCTGAGAACAG 618
 Db 655 ATCCCTTCAGAGAACAG 672

RESULT 13

BC058083

LOCUS

DEFINITION Mus musculus sodium channel, voltage-gated, type III, beta, mRNA (cDNA clone IMAGE:6826414), containing frame-shift errors.

ACCESSION BC058083

VERSION BC058083.1 GI:34784518

KEYWORDS HTC.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 4105)

AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buettow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marudina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Tishiyuki, S., Chencini, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywiecki, M.I., Skalska, U., Smalish, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

MEDLINE 22388257

PUBMED 12477932

REFERENCE 2 (bases 1 to 4105)

AUTHORS Strausberg, R.

TITLE Direct Submission

JOURNAL Submitted (08-SEP-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk
Email: cgabs-t@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: M. Bento Soares, University of Iowa
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smalish, Michael Smith, Lorraine Spence, Jeff Scott, Michael Thorne, Miranada Tsai, Natsaja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
Series: IRAK Plate: 126 Row: 1 Column: 19
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA yi: 23943798
This clone has the following problem: frame shifted.

FEATURES
Location/Qualifiers
1. .4105
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6826414"
/tissue_type="Brain, mouse, 13.5,14.5,16.5,17.5 dpc"
/clone_lib="NIH_BMAP_FY0"

ORIGIN
Query Match 78.4%; Score 508.2; DB 3; Length 4105;
Best Local Similarity 88.7%; Pred. No. 7.6e-126;
Matches 574; Conservative 0; Mismatches 68; Indels 5; Gaps 2;
/lab_host="DH10B"
/note="Vector: pYX-ASC"

QY 1 ATGCTGCTCCCTCAATAGATTGTTCCCTCGGCTCTCTCGTCTTACTTACTGGTCAGT 60
DB 356 ATGCTGCTCCCTCAACAGATTGTTCCCTCGGCTCTCTCTAGTCTCATCTACTGG---GA 411
QY 61 GTCTGCTTCCCTGTGTGTGGAGTGCCTCCGAGACGAGGCGCGTGCAGGCGCAACCCC 120
DB 412 GTCTGCTTCCCTGTGTGTGGAGTGCCTCCGAGACGAGGCGCGTGCAGGCGCAATTC 471
QY 121 ATGAAGCTGCTGCTGCTCTCTCTGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
DB 472 ATGAAGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 531
QY 181 GAATGCTTCTACAGGCGCGGCGGCTTAAAGATTTCCTTATTTACGATATCGGAATGGC 240
DB 532 GAGTGGTCTTACAGGCGCGGCGGCTTAAAGATTTCCTTATTTACGATATCGGAATGGC 591
QY 241 CACAGAGAGTGGAG 300
DB 592 CACAGAGAGTGGAG 650
QY 301 CAGGAGGTGTCATCATCTGTGCTCAACGATCATCTCTGAAAGACTCTGCGCTCTACATGC 360
DB 651 CAGGAGGTATCATCATCTGTGCTCAACGATCATCTCTGAAAGACTCTGCGCTCTACATGC 710
QY 361 AATGTCCTCCGCGAGTTTGAGTTTGAGGCGCATCGCCCTTTGTGAGAGAGAGCGCGCTG 420
DB 711 AATGTCCTCCGCGAGTTTGAGTTTGAGGCGCATCGCCCTTTGTGAGAGAGAGCGCGCTG 770
QY 421 ATCCCTCTTAAGAGTCAACAGAGAGAGTGGAGAGAGAGTTCACCTCTCTGTGGTCTCAGAAATC 480
DB 771 ATACCTCTCGAGTCACTGAGAGAGAGAGAGAGAGAGTTCACCTCTCGTGGTCTCGGAAATC 830
QY 481 ATGATGTACATCTCTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 540
DB 831 ATGATGTACATCTCTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 890
QY 541 TACAGAAAGGTCTCAAAAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
DB 891 TACAGAAAGGTCTCAAAAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 950
QY 601 ATCCCTCTGAGAACAGAGAGAGAGTCTGCGGTACCAAGTGGAGGAATA 647
DB 951 ATCCCTCTGAGAACAGAGAGAGAGTCTGCGGTACCAAGTGGAGGAATA 997

RESULT 14
BG294174
LOCUS
DEFINITION 950 bp mRNA linear EST 21-FEB-2001
602391245F1 NTH_MGC_94 Mus musculus cDNA clone IMAGE:4503250 5',
mRNA sequence.

ACCESSION BG294174
VERSION BG294174.1 GI:13054543
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 950)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLAM10373 row: i column: 11
High quality sequence stop: 643.
FEATURES
source
1. .950
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:4503250"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 94"
/note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 3.3 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."
ORIGIN
Query Match 72.2%; Score 468; DB 4; Length 950;
Best Local Similarity 88.8%; Pred. No. 3.8e-115; Indels 2; Gaps 2;
Matches 529; Conservative 0; Mismatches 65; Indels 2; Gaps 2;
QY 1 ATGCCTGCTTCAATAGATTGTTCCCTGGCTTCTCTGCTTATCTACTGGGTCACT 60
DB 87 ATGCCTGCTTCAACAGATTGTTCCCTAGCTTCTAGTGTCTATCTACTGGGTCA 146
QY 61 GTCTGCTTCCCTGTGTGTGGAGTGCCTCGGAGACGGCCGTGAGGGCAACCC 120
DB 147 GTCTGCTTCCCTGTGTGTAGAGTACCTCGGAGACAGCCGTGAGGGCAATTC 206
QY 121 ATGAAGCTGCTGATCTCTGATGAAGAGAGAGAGGTGGAGGCCACCACTGGT 180
DB 207 ATGAAGCTGATGATCTCTGATGAAGAGAGAGAGGTGGAGGCCACCACTG 266
QY 181 GAATGTTCTACAGGCCCGAGGGCGGTAAAGATTTCCTTATTACAGATATCG 240
DB 267 GAGTGGTTCTACAGGCTGAGGGCGGTAAAGATTTCCTTATTATAGTATCG 326
QY 241 CACGAGGTGGAGAGCCCTTTACGGGCGCTGACGTGAATGGCAGGACCTG 300
DB 327 CACGAGGTGGAGAGCCCTTTCAAGGTGCTGTCAGTGAATGGGCAAGACCT 386
QY 301 CAGGAGTGTCTACACTGTCTCAACGTCACTCTGAACGACTCTGGCCCTTAC 360
DB 387 CAGGAGTATCATCACTGTTCTCATGTGCTCACTGAATGACTCTGGCCCT 446
QY 361 AATGTCTCCGGAGTTTGAGTTTGAAGCGCATTCGGCCCTTTGTGAAGACG 420
DB 447 AATGTGT-CCAGGAGTTTGAAGCGACACACCGGCCCTTTGTGAAGACCA 505
QY 421 ATCCCTTAAGAGTACCGAGGCTGAGAGGACTTCACCTCTGTGCTTCAGAA 480
DB 506 ATACCTCTGCGAGTCACTGAAGAGGCGGAGAGACTTCACCTCGGTGCTCG 565
QY 481 ATGATGTACATCTCTGCTTCTTCTCACCCTGTGCTGCTCATCGAGATGAT 540
DB 566 ATGATGTACATCTCTGCTTCTTCTCCTCACTGTGCTGCTTTATTGAGAT 625
QY 541 TACAGAAAGTCTCAAAAGCCGAGAGGAGGCCCAAGAAACCGCTGCTGACT 596
DB 626 TACAGAAAGTCTCTAAGGCCGAGAGGAG-TCAGGAAATGCGTCTGACTAC 680
RESULT 15
AY419146 584 bp DNA linear GSS 12-DEC-2003
LOCUS
DEFINITION Pan troglodytes HCM6793 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.

ACCESSION AY419146
VERSION AY419146.1 GI:39775106
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Pan.
REFERENCE 1 (bases 1 to 584)
AUTHORS Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarawal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 584)
AUTHORS Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarawal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT These sequences were made by sequencing genomic exons and ordering them based on alignment.
FEATURES
source
1. .584
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>584
/locus_tag="HCM6793"
ORIGIN
Query Match 71.8%; Score 465; DB 9; Length 584;
Best Local Similarity 80.1%; Pred. No. 2.2e-114;
Matches 468; Conservative 0; Mismatches 116; Indels 0; Gaps 0;
QY 1 ATGCCTGCTTCAATAGATTGTTCCCTGGCTTCTCTGCTTATCTACTGGGTCACT 60
DB 1 ATGCCTGCTTCAATAGATTGTTCCCTGGCTTCTCTGCTTATCTACTGGGTCACT 60
QY 61 GTCTGCTTCCCTGTGTGTGGAGTGCCTCGGAGACGGCCGTGAGGGCAACCC 120
DB 61 GTCTGCTTCCCTGTGTGTGGAGTGCCTCGGAGACGGCCGTGAGGGCAACCC 120
QY 121 ATGAAGCTGCTGCTGATCTCTGATGAAGAGAGAGGTGGAGGCCACCACTGGT 180
DB 121 ATGAAGCTGCTGCTGATCTCTGATGAAGAGAGAGGTGGAGGCCACCACTGGT 180
QY 181 GAATGTTCTACAGGCCCGAGGGCGGTAAAGATTTCCTTATTACAGATATCG 240
DB 181 GAATGTTNNNNNAGGCGCTGAGGGCGGTAAAGATTNNNNNNNNNNNNNNNN 240
QY 241 CACGAGAGGTGGAGAGCCCTTTTACGGGCGCTGAGTGAATGGCAGGACCTG 300
DB 241 NNN 300
QY 301 CAGGAGCTGTCCATCACTGTGCTCAACGTCACTCTGAACGACTCTGGCCCTT 360
DB 301 CAGGAGCTGTCCATCACTGTGCTCAACGTCACTCTGAACGACTCTNNNNN 360
QY 361 AATGTCTCCGGAGTTTGAGTTTGAAGCGCATTCGGCCCTTTGTGAAGACG 420
DB 361 AATGTGTCTNNNGGANTNNNAGTTTGGAGCGCATCGGCCCTTTGTGAANA 420
QY 421 ATCCCTTAAGAGTACCGAGGCTGAGAGGACTTCACCTCTGTGCTTCAGAA 480
DB 421 ATCCCTTAAGANTNACCGAGGCTGAGAGGACTTTTANNCTGTGTTCTCA 480
QY 481 ATGATGTACATCTTCTGCTTCTTCTCACCCTGTGCTGCTCATCGAGATGAT 540

Db 481 ATGATGTACATCCTTCTGGTCTTCCTCNCCTTGTGGCTTCTCATCGAGATGATATATTGC 540
Qy 541 TACAGAAAGGTCTCAAAAGCCGAGAGGGCAGCCCAAGAAACGC 584
Db 541 NACAGAAAGGTCTCAAAAGCCGAGAGGGCAGCCCAAGAAACGC 584

Search completed: April 7, 2005, 14:55:40
Job time : 3871.9 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2005, 05:42:10 ; Search time 5458 Seconds
(without alignments)
11194.948 Million cell updates/sec

Title: US-09-977-579-4
Perfect score: 1261
Sequence: 1 cctcctctccgagctgagc.....tgccagaactgagaagccgg 1261

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 10
Total number of hits satisfying chosen parameters: 3827947

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_on.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------------|-------------|
| 1 | 1261 | 100.0 | 1261 | 6 AR359850 | Sequence |
| 2 | 1261 | 100.0 | 1261 | 6 AX039100 | Sequence |
| 3 | 1261 | 100.0 | 1261 | 9 HSA243396 | Sequence |
| 4 | 860 | 68.2 | 5306 | 9 AB032984 | Homo sapi |
| 5 | 543 | 43.1 | 606 | 6 CQ728741 | Sequence |
| 6 | 428 | 33.9 | 127347 | 2 AP000682 | Homo sapi |
| 7 | 428 | 33.9 | 144833 | 2 AC063921 | Homo sapi |
| 8 | 428 | 33.9 | 149800 | 2 AC021981 | Homo sapi |
| 9 | 428 | 33.9 | 172546 | 2 AC024604 | Homo sapi |
| 10 | 428 | 33.9 | 178169 | 9 AP002765 | Homo sapi |
| 11 | 428 | 33.9 | 181471 | 9 AC068539 | Homo sapi |
| 12 | 324 | 25.7 | 471 | 6 BD059018 | Secreted |
| 13 | 262 | 20.8 | 178431 | 2 AP002749 | Homo sapi |
| 14 | 226 | 17.9 | 172546 | 2 AC024604 | Homo sapi |
| 15 | 220 | 17.4 | 3296 | 9 AB097521 | Macaca fa |
| 16 | 218 | 17.3 | 144833 | 2 AC063921 | Homo sapi |
| 17 | 218 | 17.3 | 149800 | 2 AC021981 | Homo sapi |
| 18 | 216 | 17.1 | 178431 | 2 AP002749 | Homo sapi |
| 19 | 143 | 11.3 | 66980 | 2 AC013796 | Homo sapi |

| | | | | | | |
|------|-----|------|--------|----|-----------|--------------------|
| c 20 | 129 | 10.2 | 66980 | 2 | AC013796 | AC013796 Homo sapi |
| c 21 | 96 | 7.6 | 174285 | 2 | AC145484 | AC145484 Lemur cat |
| 22 | 38 | 3.0 | 645 | 6 | AX048004 | AX048004 Sequence |
| 23 | 38 | 3.0 | 2220 | 6 | AR359849 | AR359849 Sequence |
| 24 | 38 | 3.0 | 2220 | 6 | AX039099 | AX039099 Sequence |
| 25 | 38 | 3.0 | 2220 | 10 | RNO243395 | AJ243395 Rattus no |
| 26 | 38 | 3.0 | 2632 | 6 | AX048005 | AX048005 Sequence |
| 27 | 38 | 3.0 | 3107 | 10 | AF378093 | AF378093 Rattus no |
| 28 | 38 | 3.0 | 3108 | 6 | AX047984 | AX047984 Sequence |
| 29 | 29 | 2.3 | 670 | 10 | AY049036 | AY049036 Mus muscu |
| 30 | 29 | 2.3 | 4025 | 10 | AKI73115 | AKI73115 Mus muscu |
| 31 | 29 | 2.3 | 4169 | 10 | BC053919 | BC053919 Mus muscu |
| 32 | 29 | 2.3 | 4176 | 10 | BC058636 | BC058636 Mus muscu |
| 33 | 29 | 2.3 | 140307 | 2 | AC135353 | AC135353 Mus muscu |
| 34 | 29 | 2.3 | 238861 | 2 | AC148331 | AC148331 Mus muscu |
| 35 | 26 | 2.1 | 3910 | 10 | BC070899 | BC070899 Rattus no |
| 36 | 26 | 2.1 | 214306 | 2 | AC128723 | AC128723 Rattus no |
| 37 | 24 | 1.9 | 158801 | 9 | AC079118 | AC079118 Homo sapi |
| 38 | 24 | 1.9 | 205761 | 2 | CR558302 | CR558302 Danio rer |
| 39 | 23 | 1.8 | 257 | 3 | AY009586 | AY009586 Boophilus |
| 40 | 23 | 1.8 | 912 | 5 | BX931708 | BX931708 Gallus ga |
| 41 | 23 | 1.8 | 922 | 5 | BX935227 | BX935227 Gallus ga |
| 42 | 23 | 1.8 | 84090 | 2 | AC149895 | AC149895 Xenopus t |
| 43 | 23 | 1.8 | 109446 | 10 | AL645845 | AL645845 Mouse DNA |
| 44 | 23 | 1.8 | 173159 | 2 | BX537251 | BX537251 Danio rer |
| 45 | 23 | 1.8 | 190037 | 2 | CR394567 | CR394567 Danio rer |

ALIGNMENTS

RESULT 1
LOCUS AR359850 1261 bp DNA linear PAT 17-AUG-2003
DEFINITION Sequence 4 from patent US 6593565.
ACCESSION AR359850
VERSION AR359850.1 GI:33766660
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1261)
AUTHORS Heslin,P. and Lynam,N.R.
TITLE Vehicle interior rearview mirror assembly including an accessory-containing housing
JOURNAL Patent: US 6593565-A 4 15-JUL-2003;
FEATURES Location/Qualifiers
source 1. .1261
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

| | | | | |
|-----------------------|-----------------|---|-----------|--------------|
| Query Match | 100.0%; | Score 1261; | DB 6; | Length 1261; |
| Best Local Similarity | 100.0%; | Pred. No. 0; | | |
| Matches 1261; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |
| Qy | 1 | CCCTCCCTCCGAGCTGAGCTTACCTCGGCGCAACGAGCGAGGCGCGAGTGG | 60 | |
| Db | 1 | CCCTCCCTCCGAGCTGAGCTTACCTCGGCGCAACGAGCGAGGCGCGAGTGG | 60 | |
| Qy | 61 | AAGCTGAGTTCGCGGGTGGCGGGGAGGCGGAGTGTCCGTGCTGCTGAGCGCGCGAGA | 120 | |
| Db | 61 | AAGCTGAGTTCGCGGGTGGCGGGGAGGCGGAGTGTCCGTGCTGCTGAGCGCGCGAGA | 120 | |
| Qy | 121 | GCGGCGCGAGCGGCTGATCGGCTCCCTGAACTGGGAGGTCCTAGTGGGTCGCTTAG | 180 | |
| Db | 121 | GCGGCGCGAGCGGCTGATCGGCTCCCTGAACTGGGAGGTCCTAGTGGGTCGCTTAG | 180 | |
| Qy | 181 | GGCCCAAGCCCCACCCGCTCCAAAAGTCCACAGGCTCCCGAGGACCGGTGCTCG | 240 | |
| Db | 181 | GGCCCAAGCCCCACCCGCTCCAAAAGTCCACAGGCTCCCGAGGACCGGTGCTCG | 240 | |
| Qy | 241 | GCCTTCCTTCGTCAGAAAGTCGCCCCCTCGGGGCGAGTTCGTGCCCAAGGGTTTCCTCG | 300 | |

|||||
241 GCCTTCCTTCGTCAGAAAGTCGCCCCCTGGGGCGAGTTCGTCACAAAGGGTTCTCTCG 300
QY
301 AAAGAATCTGAGAGGGCGAGTCCTTGACCGAGGAATCTCTGTGTAGCCTTGAAGC 360
Db
301 AAAGAATCTGAGAGGGCGAGTCCTTGACCGAGGAATCTCTGTGTAGCCTTGAAGC 360
QY
361 CGCAGCCCCAGAAAGTCCTGCTCAATAGATTGTTCCCTGGCTTCTCTCGTCTT 420
Db
361 CGCAGCCCCAGAAAGTCCTGCTCAATAGATTGTTCCCTGGCTTCTCTCGTCTT 420
QY
421 ATCTACTGGTCACTGCTCTCTCTGCTGTGTGTGGAGTCCCTCGAGAGGGCC 480
Db
421 ATCTACTGGTCACTGCTCTCTCTGCTGTGTGTGGAGTCCCTCGAGAGGGCC 480
QY
481 GTGAGGGCAACCCCATGAAGCTGCGCTGCATCTCTGCAATGAAGAGAGAGGTGGAG 540
Db
481 GTGAGGGCAACCCCATGAAGCTGCGCTGCATCTCTGCAATGAAGAGAGAGGTGGAG 540
QY
541 GCCACCAACCGTGTGGAATGGTCTTACAGCCCGAGGGCGGTAAAGATTTCCTTATTAC 600
Db
541 GCCACCAACCGTGTGGAATGGTCTTACAGCCCGAGGGCGGTAAAGATTTCCTTATTAC 600
QY
601 GAGTATCGGAATGCCACAGAGGTGGAGAGCCCTTTTCAGGGCGCTGAGTGGAAAT 660
Db
601 GAGTATCGGAATGCCACAGAGGTGGAGAGCCCTTTTCAGGGCGCTGAGTGGAAAT 660
QY
661 GGAGCAAGGACCTGAGGAGCTGTCCATCACTGTGCTCAACGTCACTCTGAACGACTCT 720
Db
661 GGAGCAAGGACCTGAGGAGCTGTCCATCACTGTGCTCAACGTCACTCTGAACGACTCT 720
QY
721 GGCTCTTACACTGCAATGTGTCGAGAGTTTGTGAGTGGAGCGCATGCGCCCTTTTGG 780
Db
721 GGCTCTTACACTGCAATGTGTCGAGAGTTTGTGAGTGGAGCGCATGCGCCCTTTTGG 780
QY
781 AAGACGAGCGGCTGATCCCTTAAGAGTCAACGAGGAGGCTCGAGAGGACTTCACCTCT 840
Db
781 AAGACGAGCGGCTGATCCCTTAAGAGTCAACGAGGAGGCTCGAGAGGACTTCACCTCT 840
QY
841 GTGGTCTCAGAAATCATGATGTATCATCTTCCTGCTGCTTCCACCTGCTGCTCATC 900
Db
841 GTGGTCTCAGAAATCATGATGTATCATCTTCCTGCTGCTTCCACCTGCTGCTCATC 900
QY
901 GAGATGATATATTGCTACAGAAAGTCTCAAGAGCGAGAGGAGCGCCCAAGAAACGCG 960
Db
901 GAGATGATATATTGCTACAGAAAGTCTCAAGAGCGAGAGGAGCGCCCAAGAAACGCG 960
QY
961 TCTGACTTACCTTGCCATFCCATCTGAGAACAGGAGAACTCTCGGCTTACAGTGGAGAA 1020
Db
961 TCTGACTTACCTTGCCATFCCATCTGAGAACAGGAGAACTCTCGGCTTACAGTGGAGAA 1020
QY
1021 TAGAACAGGAGCAGTGTGACATGAGTGGCTGAACACCTGAGGAGCTGACATCCCATG 1080
Db
1021 TAGAACAGGAGCAGTGTGACATGAGTGGCTGAACACCTGAGGAGCTGACATCCCATG 1080
QY
1081 TTGAGCAATGTCAATGGCATCAGAGGGGCGCCCAAGGGGCCCATCGCTTCCCTTCATGC 1140
Db
1081 TTGAGCAATGTCAATGGCATCAGAGGGGCGCCCAAGGGGCCCATCGCTTCCCTTCATGC 1140
QY
1141 ATCCATGTTCTGTTTCAATTCATTCATCATACATCCACCTGCTCTGAGCTTTCACCTCT 1200
Db
1141 ATCCATGTTCTGTTTCAATTCATTCATCATACATCCACCTGCTCTGAGCTTTCACCTCT 1200
QY
1201 GACTCCCTAACTCCATCAGACCTCTTACGCAACCAATAGACTCTGCCAGAACTGAGAGCGG 1260
Db
1201 GACTCCCTAACTCCATCAGACCTCTTACGCAACCAATAGACTCTGCCAGAACTGAGAGCGG 1260
QY
1261 G 1261
Db
1261 G 1261

AX039100
LOCUS 1261 bp DNA linear PAT 18-NOV-2000
DEFINITION Sequence 4 from Patent WO0063367.
ACCESSION AX039100
VERSION AX039100.1 GI:11229276
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Cox, P.; Dixon, A.; Jackson, A. and Morgan, K.
TITLE A novel family of beta sub-unit proteins from a voltage-gated sodi
um channel, nucleic acids encoding them and therapeutic or
diagnostic uses there of
JOURNAL Patent: WO 0063367-A 4 26-OCT-2000;
WARNER-LAMBERT COMPANY (US) ; Cambridge University Technical
Services Limited (GB)
FEATURES
source
1..1261
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 1261; DB 6; Length 1261;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCCTCCCTCCGAGCTGAGCTTACCTGGCGCAACAGAGCGAGGCGCGAGTGG 60
Db 1 CCCTCCCTCCGAGCTGAGCTTACCTGGCGCAACAGAGCGAGGCGCGAGTGG 60
QY 61 AAGCTGGAGTTCGCGGGTGGCGGAGCGGAGCTGTCCGTGTGTGTGAGCGCGCGAGA 120
Db 61 AAGCTGGAGTTCGCGGGTGGCGGAGCGGAGCTGTCCGTGTGTGTGAGCGCGCGAGA 120
QY 121 GCGGCGCGGAGCGGCTGATCGGCTCCCTGCAACTGGGGAGGTCCAGTGGGGTTCGTTAG 180
Db 121 GCGGCGCGGAGCGGCTGATCGGCTCCCTGCAACTGGGGAGGTCCAGTGGGGTTCGTTAG 180
QY 181 GCGCCAAAGCCCGCCCGCTCCAAAGCTCCAGGGGCTCCCGAGGCAACCGGTCTCG 240
Db 181 GCGCCAAAGCCCGCCCGCTCCAAAGCTCCAGGGGCTCCCGAGGCAACCGGTCTCG 240
QY 241 GCGCTTCTTCGCTCAGAAAGTGGCGGCGAGTTCGTCCCAAGGGTTTCCTCG 300
Db 241 GCGCTTCTTCGCTCAGAAAGTGGCGGCGAGTTCGTCCCAAGGGTTTCCTCG 300
QY 301 AAAGAATCTGAGAGGGCGAGTTCCTTGACCGAGGGAATCTCTGTGTAGCCTTGAAGC 360
Db 301 AAAGAATCTGAGAGGGCGAGTTCCTTGACCGAGGGAATCTCTGTGTAGCCTTGAAGC 360
QY 361 CGCAGCCCCAGAAAGTGCCTTCAATAGATTGTTCCCTGGCTTCTCTCGTCTT 420
Db 361 CGCAGCCCCAGAAAGTGCCTTCAATAGATTGTTCCCTGGCTTCTCTCGTCTT 420
QY 421 ATCTACTGGTCACTGCTGTCTTCCCTGTGTGTGGAAAGTCCCTCGAGAGCGGAGCC 480
Db 421 ATCTACTGGTCACTGCTGTCTTCCCTGTGTGTGGAAAGTCCCTCGAGAGCGGAGCC 480
QY 481 GTGAGGGCAACCCCATGAGCTGCGCTGCATCTCTGCAATGAAGAGAGAGGTGGAG 540
Db 481 GTGAGGGCAACCCCATGAGCTGCGCTGCATCTCTGCAATGAAGAGAGAGGTGGAG 540
QY 541 GCCACCAACCGTGTGGAATGGTCTTACAGGGCGCGGCGGTAAAGATTTCCTTATTAC 600
Db 541 GCCACCAACCGTGTGGAATGGTCTTACAGGGCGCGGCGGTAAAGATTTCCTTATTAC 600
QY 601 GAGTATCGGAATGGCCACAGAGGTGGAGAGCCCTTTTCAGGGCGCTGAGTGGAAAT 660
Db 601 GAGTATCGGAATGGCCACAGAGGTGGAGAGCCCTTTTCAGGGCGCTGAGTGGAAAT 660

| | | | |
|--|------|--|------|
| Qy | 661 | GGCAGCAAGGACCTGCGAGGAGCTGTCATCACTGTGCTCAACGTCACCTCTGAACGACTCT | 720 |
| Db | 661 | GGCAGCAAGGACCTGCGAGGAGCTGTCATCACTGTGCTCAACGTCACCTCTGAACGACTCT | 720 |
| Qy | 721 | GGCTCTACACCTGCAATGTGTCCCGGAGTTTCAGTTTGGAGGCGCATCGGCCCTTTGTG | 780 |
| Db | 721 | GGCTCTACACCTGCAATGTGTCCCGGAGTTTCAGTTTGGAGGCGCATCGGCCCTTTGTG | 780 |
| Qy | 781 | AAGACGACGGGCTGTATCCCTTAAGAGTCAACGAGGAGGCTGGAGAGGACTTCACCTCT | 840 |
| Db | 781 | AAGACGACGGGCTGTATCCCTTAAGAGTCAACGAGGAGGCTGGAGAGGACTTCACCTCT | 840 |
| Qy | 841 | GTGTCTCAGAAATCATGATGATCATCTTCTGTGTCTTCTCACTGCTGGCTCTCATC | 900 |
| Db | 841 | GTGTCTCAGAAATCATGATGATCATCTTCTGTGTCTTCTCACTGCTGGCTCTCATC | 900 |
| Qy | 901 | GAGATGATATATTCTACAGAAAGGTCTCAAAAGCCGAAGAGGCGAGCCCAAGAAACGCG | 960 |
| Db | 901 | GAGATGATATATTCTACAGAAAGGTCTCAAAAGCCGAAGAGGCGAGCCCAAGAAACGCG | 960 |
| Qy | 961 | TCTGACTACCTTGGCATCCCATCTGAGAACAAAGGAGAACTCTGCGGTACCAAGTGAGGAA | 1020 |
| Db | 961 | TCTGACTACCTTGGCATCCCATCTGAGAACAAAGGAGAACTCTGCGGTACCAAGTGAGGAA | 1020 |
| Qy | 1021 | TAGAACAGGAGGAGTGTGACATGAGTGGCTGGAACACCTGAGGAGACTGGACATCCCATG | 1080 |
| Db | 1021 | TAGAACAGGAGGAGTGTGACATGAGTGGCTGGAACACCTGAGGAGACTGGACATCCCATG | 1080 |
| Qy | 1081 | TTGAGCAATCTCAATGCGATCAGGAGGCGGCCCAAGGCGCCCATCGCTTCCCTTCATGC | 1140 |
| Db | 1081 | TTGAGCAATCTCAATGCGATCAGGAGGCGGCCCAAGGCGCCCATCGCTTCCCTTCATGC | 1140 |
| Qy | 1141 | ATCCATTGTTCTGTTTCATTTCATTCATCATACATCCACCTGCGCTCTGAGCTTTCACCTCT | 1200 |
| Db | 1141 | ATCCATTGTTCTGTTTCATTTCATTCATCATACATCCACCTGCGCTCTGAGCTTTCACCTCT | 1200 |
| Qy | 1201 | GACTCCCTAATCTCATCAGACCTCTAGCGACCAATAAGACTCTGCGCAGAACTGAGAGCGG | 1260 |
| Db | 1201 | GACTCCCTAATCTCATCAGACCTCTAGCGACCAATAAGACTCTGCGCAGAACTGAGAGCGG | 1260 |
| Qy | 1261 | G 1261 | |
| Db | 1261 | G 1261 | |
| RESULT 3 | | | |
| HSA243396 | | | |
| LOCUS | | | |
| DEFINITION Homo sapiens mRNA for voltage-gated sodium channel beta-3 subunit (scn3b gene). | | | |
| ACCESSION AJ243396 | | | |
| VERSION AJ243396.2 GI:7242612 | | | |
| KEYWORDS scn3b gene; voltage-gated sodium channel beta-3 subunit. | | | |
| SOURCE Homo sapiens | | | |
| ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | |
| REFERENCE 1 | | | |
| AUTHORS Morgan, K., Stevens, E.B., Shaw, B., Cox, P., Dixon, A.K., Lee, K., Pincock, R.D., Hughes, J., Richardson, P.J., Mizuguchi, K. and Jackson, A.P. | | | |
| TITLE beta 3: an additional auxiliary subunit of the voltage-sensitive sodium channel that modulates channel gating with distinct kinetics | | | |
| JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (5), 2308-2313 (2000) | | | |
| MEDLINE 20160948 | | | |
| PUBMED 10688874 | | | |
| REFERENCE 2 | | | |
| AUTHORS Morgan, K. | | | |
| Direct Submission | | | |
| TITLE Submitted (28-JUN-1999) Morgan K., Biochemistry, University of Cambridge, Tennis Court Road, Cambridge, CB2 1QW, UNITED KINGDOM | | | |
| JOURNAL | | | |
| REMARK Revised by [4] | | | |
| REFERENCE 3 (bases 1 to 1261) | | | |

AUTHORS

Direct Submission

JOURNAL

COMMENT

FEATURES

source

1. .1261

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/tissue_type="brain"

1. .1261

/gene="scn3b"

<1. .375

/gene="scn3b"

376. .1024

/gene="scn3b"

/function="modulates channel gating kinetics"

/evidence="experimental"

/product="voltage-gated sodium channel beta-3 subunit"

/protein_id="CA876825.1"

/db_xref="GI:7160975"

/db_xref="GOA:Q9NY72"

/translation="MPANRFLPLASLILYVYVSVCPVCEVSETEAVQGNPKLRL

CISMKRESEVATTVVWFVYRPGGKDFLIYEVNRGHORVESPFQGRLOWNGSKDLDR

VSITLVNLTNDLSGLYTCNVSREFEFEARHPFKVTRLIPLRVEAGEDPTSVVSI

MMYILLVFLUTLWLLIEMIYCYRKVSKAEBAQENADYLAIPSENKENSAPVEE"

1025. .>1261

/gene="scn3b"

3'UTR

ORIGIN

Query Match 100.0%; Score 1261; DB 9; Length 1261;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCTCCCTCCCGAGCTGAGCTTACCTCGGCGCAAAACGACGAGGCGGCGGAGTGG 60

Db 1 CCCTCCCTCCCGAGCTGAGCTTACCTCGGCGCAAAACGACGAGGCGGCGGAGTGG 60

Qy 61 AAGCTGAGTTCCGGGGTGGCGGGGAGGCGAGTGTCCGTGGTGTGTGAGCGCGCGCAGA 120

Db 61 AAGCTGAGTTCCGGGGTGGCGGGGAGGCGAGTGTCCGTGGTGTGTGAGCGCGCGCAGA 120

Qy 121 GCGGGCGCGAGCGGCTGATCGGCTCCCTCGAACTGGGGAGTCCAGTGGGGTGGCTTAG 180

Db 121 GCGGGCGCGAGCGGCTGATCGGCTCCCTCGAACTGGGGAGTCCAGTGGGGTGGCTTAG 180

Qy 181 GCGCCAAAGCCCGCCACCCGCTCCAAAAGCTCCCGAGGCGCTCCCGAGGCAACCGGTGCTCG 240

Db 181 GCGCCAAAGCCCGCCACCCGCTCCAAAAGCTCCCGAGGCGCTCCCGAGGCAACCGGTGCTCG 240

Qy 241 GCCCTTCTTCGTGTCAGAAAGTCCGCCCTCGGGGCGAGTTGTCGCCAAAGGGTTTCCTCG 300

Db 241 GCCCTTCTTCGTGTCAGAAAGTCCGCCCTCGGGGCGAGTTGTCGCCAAAGGGTTTCCTCG 300

Qy 301 AAGAATCTGAGAGGCGGAGTCCCTTGACCGAGGGAATCTCTGTGTAGCCTTGGAGC 360

Db 301 AAGAATCTGAGAGGCGGAGTCCCTTGACCGAGGGAATCTCTGTGTAGCCTTGGAGC 360

Qy 361 CGCAGCGCCCGAAGATGCTGCTTCAATAGATTTTCCCTCGGCTTCTCTCGTCTT 420

Db 361 CGCAGCGCCCGAAGATGCTGCTTCAATAGATTTTCCCTCGGCTTCTCTCGTCTT 420

Qy 421 ATCTACTGGGTGAGTGTCTGCTTCCCTGTGTGTGTGAGAGTGCCTCGGAGAGCGAGGCC 480

Db 421 ATCTACTGGGTGAGTGTCTGCTTCCCTGTGTGTGTGAGAGTGCCTCGGAGAGCGAGGCC 480

Qy 481 GTGAGGGCAACCCCATGAGTCCGCTCATCTCTGATGAAGAGAGGAGGAGTGGAG 540

Db 481 GTGAGGGCAACCCCATGAGTCCGCTCATCTCTGATGAAGAGAGGAGGAGTGGAG 540

Morgan, K.

Submitted (13-MAR-2000) Morgan K., Biochemistry, University of

Cambridge, Tennis Court Road, Cambridge, CB2 1QW, UNITED KINGDOM

On Mar 14, 2000 this sequence version replaced gi:7160974.

Location/Qualifiers

1. .1261

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/tissue_type="brain"

1. .1261

/gene="scn3b"

<1. .375

/gene="scn3b"

376. .1024

/gene="scn3b"

/function="modulates channel gating kinetics"

/evidence="experimental"

/product="voltage-gated sodium channel beta-3 subunit"

/protein_id="CA876825.1"

/db_xref="GI:7160975"

/db_xref="GOA:Q9NY72"

/translation="MPANRFLPLASLILYVYVSVCPVCEVSETEAVQGNPKLRL

CISMKRESEVATTVVWFVYRPGGKDFLIYEVNRGHORVESPFQGRLOWNGSKDLDR

VSITLVNLTNDLSGLYTCNVSREFEFEARHPFKVTRLIPLRVEAGEDPTSVVSI

MMYILLVFLUTLWLLIEMIYCYRKVSKAEBAQENADYLAIPSENKENSAPVEE"

1025. .>1261

/gene="scn3b"

3'UTR

ORIGIN

Query Match 100.0%; Score 1261; DB 9; Length 1261;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1261; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCTCCCTCCCGAGCTGAGCTTACCTCGGCGCAAAACGACGAGGCGGCGGAGTGG 60

Db 1 CCCTCCCTCCCGAGCTGAGCTTACCTCGGCGCAAAACGACGAGGCGGCGGAGTGG 60

Qy 61 AAGCTGAGTTCCGGGGTGGCGGGGAGGCGAGTGTCCGTGGTGTGTGAGCGCGCGCAGA 120

Db 61 AAGCTGAGTTCCGGGGTGGCGGGGAGGCGAGTGTCCGTGGTGTGTGAGCGCGCGCAGA 120

Qy 121 GCGGGCGCGAGCGGCTGATCGGCTCCCTCGAACTGGGGAGTCCAGTGGGGTGGCTTAG 180

Db 121 GCGGGCGCGAGCGGCTGATCGGCTCCCTCGAACTGGGGAGTCCAGTGGGGTGGCTTAG 180

Qy 181 GCGCCAAAGCCCGCCACCCGCTCCAAAAGCTCCCGAGGCGCTCCCGAGGCAACCGGTGCTCG 240

Db 181 GCGCCAAAGCCCGCCACCCGCTCCAAAAGCTCCCGAGGCGCTCCCGAGGCAACCGGTGCTCG 240

Qy 241 GCCCTTCTTCGTGTCAGAAAGTCCGCCCTCGGGGCGAGTTGTCGCCAAAGGGTTTCCTCG 300

Db 241 GCCCTTCTTCGTGTCAGAAAGTCCGCCCTCGGGGCGAGTTGTCGCCAAAGGGTTTCCTCG 300

Qy 301 AAGAATCTGAGAGGCGGAGTCCCTTGACCGAGGGAATCTCTGTGTAGCCTTGGAGC 360

Db 301 AAGAATCTGAGAGGCGGAGTCCCTTGACCGAGGGAATCTCTGTGTAGCCTTGGAGC 360

Qy 361 CGCAGCGCCCGAAGATGCTGCTTCAATAGATTTTCCCTCGGCTTCTCTCGTCTT 420

Db 361 CGCAGCGCCCGAAGATGCTGCTTCAATAGATTTTCCCTCGGCTTCTCTCGTCTT 420

Qy 421 ATCTACTGGGTGAGTGTCTGCTTCCCTGTGTGTGTGAGAGTGCCTCGGAGAGCGAGGCC 480

Db 421 ATCTACTGGGTGAGTGTCTGCTTCCCTGTGTGTGTGAGAGTGCCTCGGAGAGCGAGGCC 480

Qy 481 GTGAGGGCAACCCCATGAGTCCGCTCATCTCTGATGAAGAGAGGAGGAGTGGAG 540

Db 481 GTGAGGGCAACCCCATGAGTCCGCTCATCTCTGATGAAGAGAGGAGGAGTGGAG 540


```

Qy 889 TGGCTGCTCATCGAGATGATATATTGCTACAGAAAGGTCTCAAGAGCGTCCGAGGCGAGCC 948
Db 561 TGGCTGCTCATCGAGATGATATATTGCTACAGAAAGGTCTCAAGAGCGTCCGAGGCGAGCC 620
Qy 949 CAAGAAAACGGTCTGACTACCTTGGCCATCCCATCTGAGAACAAAGAGAACTCTGCGGTA 1008
Db 621 CAAGAAAACGGTCTGACTACCTTGGCCATCCCATCTGAGAACAAAGAGAACTCTGCGGTA 680
Qy 1009 CCAGTGGAGGAATAGAACAGGAGCAGTGTGACATGAGTGGCTGACACCTGAGGAGCT 1068
Db 681 CCAGTGGAGGAATAGAACAGGAGCAGTGTGACATGAGTGGCTGACACCTGAGGAGCT 740
Qy 1069 GGACATCCCATGTTTCAGCAATGTCAATGGCATCAGGAGGCGCCCAAGGCGCCCATCGC 1128
Db 741 GGACATCCCATGTTTCAGCAATGTCAATGGCATCAGGAGGCGCCCAAGGCGCCCATCGC 800
Qy 1129 TTCCCTTCATGCATCCATTTGTTTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 1188
Db 801 TTCCCTTCATGCATCCATTTGTTTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 860
Qy 1189 GCTTTACCTCTGACTCCCTAACTCCATCAGACCTCTACGCACCATTAAGACTCTGCCAGA 1248
Db 861 GCTTTACCTCTGACTCCCTAACTCCATCAGACCTCTACGCACCATTAAGACTCTGCCAGA 920
Qy 1249 ACTGAGAAGCC 1259
Db 921 ACTGAGAAGCC 931

RESULT 5
LOCUS CQ728741 606 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 14675 from Patent WO02068579.
ACCESSION CQ728741
VERSION CQ728741.1 GI:42298386
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
FEATURES
source KITS, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
Patent: WO 02068579-A 14675 06-SEP-2002;
PE Corporation (NY) (US)
Location/Qualifiers
source 1..606
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 43.1%; Score 543; DB 6; Length 606;
Best Local Similarity 99.8%; Pred. No. 1.7e-299;
Matches 593; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 430 GTGAGTGTCTGCTCCCTGCTGTGTGTGAGAGTGCCCTCGAGAGCGAGGCGGTCGAGGC 489
Db 13 GTGAGTGTCTGCTCCCTGCTGTGTGTGAGAGTGCCCTCGAGAGCGAGGCGGTCGAGGC 72
Qy 490 AACCCCATGAAGCTGCGCTGCTGCTGCTGATGAAGAGAGAGGAGGAGGCGGCGGCGGCG 549
Db 73 AACCCCATGAAGCTGCGCTGCTGCTGCTGATGAAGAGAGAGGAGGAGGCGGCGGCGGCG 132
Qy 550 GTGTGGAATGTTTCTACAGCCCGGAGGCGGTTAAAGATTTCCTTATTACGAGTATCGG 609
Db 133 GTGTGGAATGTTTCTACAGCCCGGAGGCGGTTAAAGATTTCCTTATTACGAGTATCGG 192
Qy 610 AATGCCACACGAGGTGGAGAGCCCTTTTCAGGGGCGCTGCAAGTGGATGGCAGCAAG 669
Db 193 AATGCCACACGAGGTGGAGAGCCCTTTTCAGGGGCGCTGCAAGTGGATGGCAGCAAG 252

```

```

Qy 670 GACCTGCAGGACGCTGTCCATCACTGTGCTCAACGTCACCTCTGAACGACTCTGGCCTCTAC 729
Db 253 GACCTGCAGGACGCTGTCCATCACTGTGCTCAACGTCACCTCTGAACGACTCTGGCCTCTAC 312
Qy 730 ACCTGCAATGTGTCCCGGAGTTTGTAGTTTGTAGGCGCATCGGCCCTTTTGTGAAGACGACG 789
Db 313 ACCTGCAATGTGTCCCGGAGTTTGTAGTTTGTAGGCGCATCGGCCCTTTTGTGAAGACGACG 372
Qy 790 CGGCTGATCCCTTAAGAGTCACCGAGGAGGCTGGAGAGGACTTCACTCTGTGGTCTCA 849
Db 373 CGGCTGATCCCTTAAGAGTCACCGAGGAGGCTGGAGAGGACTTCACTCTGTGGTCTCA 432
Qy 850 GAAATCATGATGTACATCTTCTGCTCTTCTCTCACTCTGCTGCTCATCGAGATGATA 909
Db 433 GAAATCATGATGTACATCTTCTGCTCTTCTCTCACTCTGCTGCTCATCGAGATGATA 492
Qy 910 TATTGTACAGAAAGTCTCAAAAGCGAAGAGGCGCCCAAGAAACCGCTGTGACTAC 969
Db 493 TATTGTACAGAAAGTCTCAAAAGCGAAGAGGCGCCCAAGAAACCGCTGTGACTAC 552
Qy 970 CTTGCCATCCCATCTGAGACAGGAGACTCTGCGGTACCACTGAGGGAATAG 1023
Db 553 CTTGCCATCCCATCTGAGAAACAAGSAGAACTCTGCGGTACCACTGAGGGAATAG 606

RESULT 6
AP000682/c 127347 bp DNA linear HTG 16-AUG-2000
LOCUS Homo sapiens chromosome 11 clone CMB9-32A1 map 11q24, WORKING DRAFT
DEFINITION SEQUENCE, 8 unordered pieces.
ACCESSION AP000682
VERSION AP000682.3 GI:9844967
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,
AUTHORS Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE Homo sapiens 127,347 genomic DNA of 11q24
JOURNAL Published Only in DataBase (1999)
REFERENCE 2
AUTHORS Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE Direct Submission
JOURNAL Submitted (08-NOV-1999) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555, Japan
[8-mail:hattori@gsc.riken.go.jp, URL:http://bgp.gsc.riken.go.jp/,
Tel:81-42-778-9923, Fax:81-42-778-9924]
COMMENT On Aug 18, 2000 this sequence version replaced gi:8118870.
----- Genome Center
Center: RIKEN Genomic Sciences Center (GSC)
Center code: RIKEN
Web site: http://bgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
----- Project Information
Center project name: HumDrafl1
Center clone name: CMB9-32A1
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 124981 bases at least Q40
Consensus quality: 125815 bases at least Q30
Consensus quality: 126286 bases at least Q20
Insert size: 126647; sum-of-contigs
Quality coverage: 12.71x in Q20 bases; sum-of-contigs
-----
NOTE: This is a 'working draft' sequence. It currently consists of
8 contigs. The true order of the pieces is not known and their

```

order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

```

1      33532 contig of 33532 bp in length
33533      59223 contig of 25591 bp in length
59224      83007 contig of 23684 bp in length
83108      100196 contig of 17089 bp in length
100297      112313 contig of 12017 bp in length
112414      121043 contig of 8630 bp in length
121144      125917 contig of 4774 bp in length
126018      127347 contig of 1330 bp in length.

```

* NOTE: This is a 'working draft' sequence. It currently consists of 8 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```

1      33532: contig of 33532 bp in length
*      33533: gap of 100 bp
*      59223: contig of 25591 bp in length
*      59224: gap of 100 bp
*      83007: contig of 23684 bp in length
*      83108: gap of 100 bp
*      100196: contig of 17089 bp in length
*      100297: gap of 100 bp
*      112313: contig of 12017 bp in length
*      112314: gap of 100 bp
*      121043: contig of 8630 bp in length
*      121044: gap of 100 bp
*      121144: contig of 4774 bp in length
*      125917: gap of 100 bp
*      126018: contig of 1330 bp in length.

```

FEATURES

```

source
1. .127347
   /organism="Homo sapiens"
   /mol_type="genomic DNA"
   /db_xref="taxon:9606"
   /chromosomes="11"
   /map="11q24"
   /clone="CMB9-32A1"

misc_feature
1. .33532
   /note="assembly_fragment"
33533. .59223
   /note="assembly_fragment"
59224. .83007
   /note="assembly_fragment"
83108. .100196
   /note="assembly_fragment clone_end:SP6 vector_side:left"
100297. .112313
   /note="assembly_fragment"
112414. .121043
   /note="assembly_fragment clone_end:T7 vector_side:right"
121144. .125917
   /note="assembly_fragment"
126018. .127347
   /note="assembly_fragment"

```

ORIGIN

```

Query Match      33.9%; Score 428; DB 2; Length 127347;
Best Local Similarity 100.0%; Pred. No. 1.9e-233;
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3      CTCCTTCGAGTGTACCTCGGCGCAACGAGCGAGCGAGGGCGGAGTGGAA 62
Db      29132  CTCCTTCGAGTGTACCTCGGCGCAACGAGCGAGCGAGGGCGGAGTGGAA 29073

QY      63      GTGGAGTTCGGGGTGGGGAGGCGAGCTGTCCGTGGTGTGTGAGCGCCGCGAGAGC 122
Db      29072  GCTGGAGTTCGGGGTGGGGAGGCGAGCTGTCCGTGGTGTGTGAGCGCCGCGAGAGC 29013

```

```

QY      123      GGGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGAGGTCCAGTGGGGTGCCTTAGGG 182
Db      29012  GGGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGAGGTCCAGTGGGGTGCCTTAGGG 28953

QY      183      CCCAAAGCCCCCACCAGCGCTCCAAAAGTCCCCAGAGGCTCCCGAGGACCGGTGCTCGGC 242
Db      28952  CCCAAAGCCCCCACCAGCGCTCCAAAAGTCCCGAGGCTCCCGAGGACCGGTGCTCGGC 28893

QY      243      CTTCTCTTCGGTTCAGAAAGTTCGCCCTCGGGGCGAGTTCGTCCCAAGAGGTTTCCTCGAA 302
Db      28892  CTTCTCTTCGGTTCAGAAAGTTCGCCCTCGGGGCGAGTTCGTCCCAAGAGGTTTCCTCGAA 28833

QY      303      AGAATCTGAGAGGGCGGAGTTCCTGACGAGGGAATCTCTGTGTAGCTTGGAGCGC 362
Db      28832  AGAATCTGAGAGGGCGGAGTTCCTGACGAGGGAATCTCTGTGTAGCTTGGAGCGC 28773

QY      363      CCAGCCCCAGAGAGTCCCTTCAATAGATTTGTTCCCTGGCTTCTCTGTGCTTAT 422
Db      28772  CCAGCCCCAGAGAGTCCCTTCAATAGATTTGTTCCCTGGCTTCTCTGTGCTTAT 28713

QY      423      CTACTGGG 430
Db      28712  CTACTGGG 28705

RESULT 7
AC063921/c
LOCUS      AC063921      144833 bp      DNA      linear      HTG 08-JAN-2003
DEFINITION Homo sapiens chromosome 3 clone RP11-142P10, WORKING DRAFT
AC063921
VERSION      AC063921.18      GI:20335587
HTG: HTGS PHASE1; HTGS_DRAFT.
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 144833)
AUTHORS      Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C.,
      Alsbrooks, S.L., Amaratunge, H.C., Are, J.R., Ayale, M., Banks, T.,
      Barbara, J., Benton, J., Bimage, K., Blankenburg, K., Bonnin, D.,
      Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,
      Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., C.,
      Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
      Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
      Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
      Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
      Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
      Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
      Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
      Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
      Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
      Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
      Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
      Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B.,
      Hombl, P., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,
      Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
      Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
      Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
      Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulaseg, H.,
      Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
      Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, B.,
      Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M.,
      Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
      Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
      Nguyen, N., Nickerson, E., Nwokenko, S., Ogih, M., Okwuonu, G.,
      Oreguene, N., Oviado, R., Pace, A., Payton, B., Peery, J., Perez, L.,
      Peters, L., Picken, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
      Rivers, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savery, G.,
      Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sisson, I.,
      Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,
      Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,

```


Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Umani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wlezyk, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.

TITLE JOURNAL REFERENCE

Unpublished
2 (bases 1 to 144833)
Worley, K.C.

TITLE JOURNAL REFERENCE

Submitted (22-APR-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 144833)
Worley, K.C.

TITLE JOURNAL REFERENCE

Submitted (08-JAN-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Apr 28, 2002 this sequence version replaced gi:16117924.

COMMENT

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: HAXY

Center clone name: RP11-142P10

----- Summary Statistics

Sequencing vector: M13;

Chemistry: Dye-terminator Big Dye; 14% of reads

Assembly: Dye-terminator Big Dye; 86% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 127260 bases at least Q40

Consensus quality: 132735 bases at least Q30

Consensus quality: 136183 bases at least Q20

Estimated insert size: 139078; sum-of-contigs estimation

Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

NOTE: This is a 'working draft' sequence. It currently

consists of 13 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

1 2437: contig of 2437 bp in length

2438 2537: gap of unknown length

2538 4812: contig of 2275 bp in length

4813 4912: gap of unknown length

4913 8488: contig of 3576 bp in length

8489 8588: gap of unknown length

8589 11163: contig of 2575 bp in length

11164 11263: gap of unknown length

11264 16638: contig of 5375 bp in length

16639 16738: gap of unknown length

16739 21870: contig of 5132 bp in length

21871 21970: gap of unknown length

21971 30780: contig of 8810 bp in length

30781 30880: gap of unknown length

30881 39267: contig of 8387 bp in length

39268 39367: gap of unknown length

39368 46890: contig of 7523 bp in length

46891 46990: gap of unknown length

46991 59656: contig of 12666 bp in length

59657 59758: gap of unknown length

59759 83309: contig of 23553 bp in length

83310 83409: gap of unknown length

83410 111857: contig of 28448 bp in length

FEATURES
source

* 111858 111957: gap of unknown length
* 111958 144833: contig of 32876 bp in length.
Location/Qualifiers
1. 144833
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="3"
/clone="RP11-142P10"

ORIGIN

Query Match 33.9%; Score 428; DB 2; Length 144833;
Best Local Similarity 100.0%; Pred. No. 1.9e-233;
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3 CTCCCTCCGAGCTGAGCTTACCTGGCGCGCAACGAGCGAGCGGGGCGGAGTGA 62
Db 50353 CTCCTTCGAGCTGAGCTTACCTGGCGCGCAACGAGCGAGCGGGGCGGAGTGA 50294
Qy 63 GCTGGAGTTCGGGGTGGCGGGGAGCGGAGCTGTCGCTGCTGAGCGCGGCGAGAGC 122
Db 50293 GCTGGAGTTCGGGGTGGCGGGGAGCGGAGCTGTCGCTGCTGAGCGCGGCGAGAGC 50234
Qy 123 GGGCGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGAGGTCAGAGTGGGTCGCTTAGGG 182
Db 50233 GGGCGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGAGGTCAGAGTGGGTCGCTTAGGG 50174
Qy 183 CCAGAGCCCCCAGCGGCTCCAGAGGTCGAGGCGGCTCCAGAGGACCGGTCGCTGGC 242
Db 50173 CCAGAGCCCCCAGCGGCTCCAGAGGTCGAGGCGGCTCCAGAGGACCGGTCGCTGGC 50114
Qy 243 CCTTCCTTCGGTCAGAAAGTCGCCCCCTGGGGGAGTTCGTCGCAAGGGTTCTCTCGAA 302
Db 50113 CCTTCCTTCGGTCAGAAAGTCGCCCCCTGGGGGAGTTCGTCGCAAGGGTTCTCTCGAA 50054
Qy 303 AGAATCTGAGAGGCGGAGCTTCCTGACCGGGAATCTCTGTGTAGCCCTTGAAGCCG 362
Db 50053 AGAATCTGAGAGGCGGAGCTTCCTGACCGGGAATCTCTGTGTAGCCCTTGAAGCCG 49994
Qy 363 CGAGCCCCAGAGATGCTGCTCCCTCAATAGATTGTTCCCTGCTCTCTGCTGCTAT 422
Db 49993 CGAGCCCCAGAGATGCTGCTCCCTCAATAGATTGTTCCCTGCTCTCTGCTGCTAT 49934
Qy 423 CTACTGGG 430
Db 49933 CTACTGGG 49926

RESULT 8

AC021981/c 149800 bp DNA linear HTG 19-JUL-2000
LOCUS
DEFINITION Homo sapiens chromosome 11 clone RP11-11C15 map 11, WORKING DRAFT
SEQUENCE, 29 unordered pieces.
AC021981
ACCESSION AC021981.2 GI:9280747
VERSION HTG; HTGS PHASE1; HTGS_DRAFT.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 149800)
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE Homo sapiens chromosome 11, clone RP11-11C15
JOURNAL Unpublished
2 (bases 1 to 149800)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, B., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Beckerly, R., Bede, F.,
Boguslavsky, L., Boukhalter, B., Brown, A., Burkett, G., Castle, A.,
Choepl, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,
Dearellano, K., Dewar, K., Domino, M., Doyle, M., Fenesor, J.,
Ferreira, P., FitzHugh, W., Forrest, C., Gage, D., Galagan, J.,
Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J.C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,

| | | | |
|---|-------|--|-------|
| misc_feature | | 43770..49948 | |
| misc_feature | | /note="assembly_fragment" | |
| misc_feature | | 50049..55578 | |
| misc_feature | | /note="assembly_fragment" | |
| misc_feature | | clone_end:SP6 | |
| misc_feature | | vector_side:right | |
| misc_feature | | 55679..60959 | |
| misc_feature | | /note="assembly_fragment" | |
| misc_feature | | 61060..66394 | |
| misc_feature | | /note="assembly_fragment" | |
| misc_feature | | 66495..72947 | |
| misc_feature | | /note="assembly_fragment" | |
| misc_feature | | 73048..80874 | |
| misc_feature | | /note="assembly_fragment" | |
| misc_feature | | 80975..87847 | |
| misc_feature | | /note="assembly_fragment" | |
| misc_feature | | 87948..97279 | |
| misc_feature | | /note="assembly_fragment" | |
| misc_feature | | 97380..106037 | |
| misc_feature | | /note="assembly_fragment" | |
| misc_feature | | 106138..114206 | |
| misc_feature | | /note="assembly_fragment" | |
| misc_feature | | 114307..129830 | |
| misc_feature | | /note="assembly_fragment" | |
| misc_feature | | 129931..149800 | |
| misc_feature | | /note="assembly_fragment" | |
| ORIGIN | | | |
| Query Match | | 33.9%; Score 428; DB 2; Length 149800; | |
| Best Local Similarity | | 100.0%; Pred. No. 1.9e-233; | |
| Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | | | |
| Qy | 3 | CTCCCTTCGAGCTGAGCTTACCTCGGCGCAACGAGCGAGCGGCGGCGCGAGTGGAA | 62 |
| Db | 19234 | CTCCCTTCGAGCTGAGCTTACCTCGGCGCAACGAGCGAGCGGCGGCGAGTGGAA | 19175 |
| Qy | 63 | GCTGGAGTTCGGGGTGGGGGAGGAGCGACTGTCGTGTGCTGAGCGCGCGAGAGC | 122 |
| Db | 19174 | GCTGGAGTTCGGGGTGGGGGAGGAGCGACTGTCGTGTGCTGAGCGCGCGAGAGC | 19115 |
| Qy | 123 | GGGGCGGAGGGGTGATCGGCTCCCTCGAACTCGGAGGTCGAGTGGGTTCGTTAGGG | 182 |
| Db | 19114 | GGGGCGGAGGGGTGATCGGCTCCCTCGAACTCGGAGGTCGAGTGGGTTCGTTAGGG | 19055 |
| Qy | 183 | CCCAAGCCCGCCACCGCTCCAAAGCTCCAGGGGCTCCCGAGGACCGGTGCTCGGC | 242 |
| Db | 19054 | CCCAAGCCCGCCACCGCTCCAAAGCTCCAGGGGCTCCCGAGGACCGGTGCTCGGC | 18995 |
| Qy | 243 | CTTCCTTCGGTCAAGAGTGGCCCTCGGGGCGAGTTCTGTCCTCAAGGGTTTCCTCGAA | 302 |
| Db | 18994 | CTTCCTTCGGTCAAGAGTGGCCCTCGGGGCGAGTTCTGTCCTCAAGGGTTTCCTCGAA | 18935 |
| Qy | 303 | AGAACTGAGAGGGCGGAGTCTTTGACCGAGGGAATCTCTGTGTAGCTTGAAGCCG | 362 |
| Db | 18934 | AGAACTGAGAGGGCGGAGTCTTTGACCGAGGGAATCTCTGTGTAGCTTGAAGCCG | 18875 |
| Qy | 363 | CCAGCCCGAGAGTGCCTGCTTCATAGATTGTTTCCCTGCTCTCTCGTGCTTAT | 422 |
| Db | 18874 | CCAGCCCGAGAGTGCCTGCTTCATAGATTGTTTCCCTGCTCTCTCGTGCTTAT | 18815 |
| Qy | 423 | CTACTGGG 430 | |
| Db | 18814 | CTACTGGG 18807 | |
| RESULT 9 | | | |
| AC024604 | | | |
| LOCUS | | | |
| DEFINITION Homo sapiens chromosome Chromosome 10 clone RP11-331G19, WORKING | | | |
| DRAFT SEQUENCE, 22 unordered pieces. | | | |
| AC024604 | | | |
| AC024604.3 GI:8389428 | | | |
| KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_CANCELLED. | | | |
| SOURCE | | | |
| ORGANISM | | | |
| Homo sapiens (human) | | | |
| Eukaryota; Metazoa; Craniata; Vertebrata; Euteleostomi; | | | |
| Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | |
| REFERENCE | | | |
| AUTHORS | | | |
| TITLE | | | |
| JOURNAL | | | |
| AUTHORS | | | |
| TITLE | | | |
| JOURNAL | | | |
| COMMENT | | | |
| On Jun 9, 2000 this sequence version replaced gi:7549605. | | | |
| Genome Center | | | |
| Center: Genome Therapeutics Corporation | | | |
| Center code: GTC | | | |
| Web site: http://www.genomecorp.com/ | | | |
| Contact: gtc-seqcenter@genomecorp.com | | | |
| ----- Project Information | | | |
| ----- Project name: hg202 | | | |
| ----- Summary Statistics | | | |
| Sequencing vector: N/A | | | |
| Chemistry: Dye-terminator Big Dye; 100% of reads | | | |
| Assembly program: Phrap; version 990315 | | | |
| Consensus quality: 152033 bases at least Q40 | | | |
| Consensus quality: 163535 bases at least Q30 | | | |
| Consensus quality: 165256 bases at least Q20 | | | |
| Insert size: 170446; sum-of-contigs | | | |
| Quality coverage: 4.3x in Q20 bases; sum-of-contigs | | | |
| ----- | | | |
| * NOTE: This is a 'working draft' sequence. It currently | | | |
| * consists of 22 contigs. The true order of the pieces | | | |
| * is not known and their order in this sequence record is | | | |
| * arbitrary. Gaps between the contigs are represented as | | | |
| * runs of N, but the exact sizes of the gaps are unknown. | | | |
| * This record will be updated with the finished sequence | | | |
| * as soon as it is available and the accession number will | | | |
| * be preserved. | | | |
| ----- | | | |
| 1 1339: contig of 1339 bp in length | | | |
| * 1340: gap of unknown length | | | |
| * 1440: contig of 1350 bp in length | | | |
| * 2790: gap of unknown length | | | |
| * 2899: contig of 1673 bp in length | | | |
| * 4562: gap of unknown length | | | |
| * 4663: contig of 1682 bp in length | | | |
| * 6345: gap of unknown length | | | |
| * 6445: contig of 2777 bp in length | | | |
| * 9221: gap of unknown length | | | |
| * 9321: contig of 2676 bp in length | | | |
| * 11997: gap of unknown length | | | |
| * 12097: contig of 2770 bp in length | | | |
| * 14867: gap of unknown length | | | |
| * 14967: contig of 4952 bp in length | | | |
| * 14968: gap of unknown length | | | |
| * 19919: contig of 4106 bp in length | | | |
| * 20019: gap of unknown length | | | |
| * 24125: contig of 5219 bp in length | | | |
| * 24225: gap of unknown length | | | |
| * 29444: contig of 5024 bp in length | | | |
| * 29544: gap of unknown length | | | |
| * 34568: contig of 4746 bp in length | | | |
| * 34668: gap of unknown length | | | |
| * 34669: contig of 4746 bp in length | | | |
| * 39415: gap of unknown length | | | |
| * 39514: contig of 6371 bp in length | | | |
| * 45885: gap of unknown length | | | |
| * 45886: contig of 5012 bp in length | | | |
| * 50997: gap of unknown length | | | |
| * 51097: contig of 8046 bp in length | | | |
| * 51998: gap of unknown length | | | |
| * 59143: contig of 8536 bp in length | | | |
| * 59244: gap of unknown length | | | |
| * 67779: contig of 8550 bp in length | | | |
| * 67780: gap of unknown length | | | |
| * 76429: contig of 8550 bp in length | | | |

* 76430 76529: gap of unknown length
* 76530 86704: contig of 10175 bp in length
* 76705 86804: gap of unknown length
* 76805 103147: contig of 16343 bp in length
* 103148 103247: gap of unknown length
* 103248 116343: contig of 13096 bp in length
* 116344 116443: gap of unknown length
* 116444 135318: contig of 18875 bp in length
* 135319 172546: gap of unknown length
* 135419 172546: contig of 37128 bp in length.

FEATURES
source
Location/Qualifiers
1. .172546
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="Chromosome 10"
/clone="RP11-331G19"
/clone_lib="RPCI-11"

ORIGIN

Query Match 33.9%; Score 428; DB 2; Length 172546;
Best Local Similarity 100.0%; Pred. No. 1.9e-233;
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CTCCTTCCGAGCTGAGCTTACCTCGGCGCAACGAGCGAGCGAGGGCGGAGTGGAA 62
Db 122377 CTCCTTCCGAGCTGAGCTTACCTCGGCGCAACGAGCGAGCGAGGGCGGAGTGGAA 122436
QY 63 GTTGGAGTTCCGGGTGGGCGGAGCGACTGTCCGTGGTCTGAGCGCGCGCGAGAGC 122
Db 122437 GCTGGAGTTCCGGGTGGGCGGAGCGACTGTCCGTGGTCTGAGCGCGCGCGAGAGC 122496
QY 123 GGGCGCGAGCGGCTGATCGGCTCCCTCGAATCTGGGAGGTCAGTGGGGTGCCTTAGGG 182
Db 122497 GGGCGCGAGCGGCTGATCGGCTCCCTCGAATCTGGGAGGTCAGTGGGGTGCCTTAGGG 122556
QY 183 CCCAAGCCCCCACCACCGGCTCCAAAGCTCCAGAGGCTCCCGAGCGACCGGTGCTCGGC 242
Db 122557 CCCAAGCCCCCACCACCGGCTCCAAAGCTCCAGAGGCTCCCGAGCGACCGGTGCTCGGC 122616
QY 243 CTTCTCTCGGTGAGAAAGTCGCCCCCTCGGGGCGAGTTCGTCCTCAAGGGTTTCTTCGAA 302
Db 122617 CTTCTCTCGGTGAGAAAGTCGCCCCCTCGGGGCGAGTTCGTCCTCAAGGGTTTCTTCGAA 122676
QY 303 AGAATCTGAGAGCGCGAGCTCTTGAACGAGGGAATCTCTGTGTAGCCTTGGAGCGCG 362
Db 122677 AGAATCTGAGAGCGCGAGCTCTTGAACGAGGGAATCTCTGTGTAGCCTTGGAGCGCG 122736
QY 363 CCAGCCCCCAGAGATGCTCCCTGCAATAGATTGTTTCCCTGGCTTCTCTGCTGCTTAT 422
Db 122737 CCAGCCCCCAGAGATGCTCCCTGCAATAGATTGTTTCCCTGGCTTCTCTGCTGCTTAT 122796
QY 423 CTACTGGG 430
Db 122797 CTACTGGG 122804

RESULT 10
AP002765/c
LOCUS AP002765 178169 bp DNA linear PRI 15-MAR-2003
DEFINITION Homo sapiens genomic DNA, chromosome 11q clone.RP11-634B22,
complete sequence.
ACCESSION AP002765
VERSION AP002765.3 GI:16751488
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Homo sapiens genomic DNA
REFERENCE
AUTHORS
TITLE

JOURNAL Published Only in Database (2000)
REFERENCE 2 (bases 1 to 178169)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (10-JUL-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suenho-chou,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@sc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/
Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT On Nov 5, 2001 this sequence version replaced gi:12381934.

FEATURES
source
Location/Qualifiers
1. .178169
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="RP11-634B22"

ORIGIN

Query Match 33.9%; Score 428; DB 9; Length 178169;
Best Local Similarity 100.0%; Pred. No. 1.9e-233;
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 CTCCTTCCGAGCTGAGCTTACCTCGGCGCAACGAGCGAGCGAGGGCGGAGTGGAA 62
Db 107533 CTCCTTCCGAGCTGAGCTTACCTCGGCGCAACGAGCGAGCGAGGGCGGAGTGGAA 107474
QY 63 GCTGGAGTTCCGGGTGGGCGGAGCGACTGTCCGTGGTCTGAGCGCGCGCGAGAGC 122
Db 107473 GCTGGAGTTCCGGGTGGGCGGAGCGACTGTCCGTGGTCTGAGCGCGCGCGAGAGC 107414
QY 123 GGGCGCGAGCGGCTGATCGGCTCCCTCGAATCTGGGAGGTCAGTGGGGTGCCTTAGGG 182
Db 107413 GGGCGCGAGCGGCTGATCGGCTCCCTCGAATCTGGGAGGTCAGTGGGGTGCCTTAGGG 107354
QY 183 CCCAAGCCCCCACCACCGGCTCCAAAGCTCCAGAGGCTCCCGAGCGACCGGTGCTCGGC 242
Db 107353 CCCAAGCCCCCACCACCGGCTCCAAAGCTCCAGAGGCTCCCGAGCGACCGGTGCTCGGC 107294
QY 243 CTTCTCTCGGTGAGAAAGTCGCCCCCTCGGGGCGAGTTCGTCCTCAAGGGTTTCTTCGAA 302
Db 107293 CTTCTCTCGGTGAGAAAGTCGCCCCCTCGGGGCGAGTTCGTCCTCAAGGGTTTCTTCGAA 107234
QY 303 AGAATCTGAGAGCGCGAGCTCTTGAACGAGGGAATCTCTGTGTAGCCTTGGAGCGCG 362
Db 107233 AGAATCTGAGAGCGCGAGCTCTTGAACGAGGGAATCTCTGTGTAGCCTTGGAGCGCG 107174
QY 363 CCAGCCCCCAGAGATGCTCCCTGCAATAGATTGTTTCCCTGGCTTCTCTGCTGCTTAT 422
Db 107173 CCAGCCCCCAGAGATGCTCCCTGCAATAGATTGTTTCCCTGGCTTCTCTGCTGCTTAT 107114
QY 423 CTACTGGG 430
Db 107113 CTACTGGG 107106

RESULT 11
AC069539/c
LOCUS AC069539 181471 bp DNA linear PRI 05-AUG-2004
DEFINITION Homo sapiens chromosome 11 clone RP11-321E15, complete sequence.
ACCESSION AC069539
VERSION AC069539.5 GI:20270093
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Homo sapiens genomic DNA
REFERENCE
AUTHORS
TITLE

RESULT 12
BD059018
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

2 (bases 1 to 178431)
 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 Direct Submission
 Submitted (10-JUL-2000) Masahira Hattori, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555, Japan
 (E-mail: hattori@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/
 Tel:81-42-778-9923, Fax:81-42-778-9924)

COMMENT

----- Genomic Center
 Center: RIKEN Genomic Sciences Center (GSC)
 Center code: RIKEN
 Web site: http://hgp.gsc.riken.go.jp/
 Contact: hattori@gsc.riken.go.jp
 ----- Project Information
 Center project name: HumDraft11
 Center clone name: RP11-158J21
 ----- Summary Statistics
 Sequencing vector: PCR products; 100% of reads
 Chemistry: Dye-terminator ET-amersham; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 155982 bases at least Q40
 Consensus quality: 165758 bases at least Q30
 Consensus quality: 170239 bases at least Q20
 Insert size: 174431; sum-of-contigs
 Quality coverage: 4.20x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of
 41 contigs. The true order of the pieces is not known and their
 order in this sequence record is arbitrary. Gaps between the
 contigs are represented as runs N, but the exact sizes of the gaps
 are unknown. This record will be updated with the finished sequence
 as soon as it is available and the accession number will be
 preserved.

1
 13012 contig of 13012 bp in length
 13113 contig of 11565 bp in length
 24677 contig of 9863 bp in length
 34640 contig of 8456 bp in length
 43196 contig of 8456 bp in length
 50126 contig of 6830 bp in length
 57725 contig of 7499 bp in length
 57826 contig of 6287 bp in length
 64612 contig of 6287 bp in length
 64713 contig of 6287 bp in length
 70999 contig of 6287 bp in length
 71100 contig of 6287 bp in length
 76188 contig of 6287 bp in length
 76289 contig of 6287 bp in length
 83681 contig of 6287 bp in length
 90511 contig of 6287 bp in length
 90612 contig of 6287 bp in length
 95494 contig of 6287 bp in length
 95594 contig of 6287 bp in length
 102363 contig of 6287 bp in length
 105967 contig of 6287 bp in length
 11142 contig of 6287 bp in length
 115923 contig of 6287 bp in length
 116024 contig of 6287 bp in length
 121747 contig of 6287 bp in length
 125332 contig of 6287 bp in length
 125633 contig of 6287 bp in length
 128028 contig of 6287 bp in length
 131184 contig of 6287 bp in length
 132826 contig of 6287 bp in length
 135667 contig of 6287 bp in length
 138001 contig of 6287 bp in length
 138101 contig of 6287 bp in length
 141194 contig of 6287 bp in length
 143619 contig of 6287 bp in length
 143720 contig of 6287 bp in length
 146448 contig of 6287 bp in length
 149531 contig of 6287 bp in length
 152157 contig of 6287 bp in length
 155624 contig of 6287 bp in length
 158335 contig of 6287 bp in length
 160510 contig of 6287 bp in length
 162343 contig of 6287 bp in length
 174281 contig of 6287 bp in length

175830 176905 contig of 1076 bp in length
 177006 178431 contig of 1426 bp in length.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 41 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1
 13012: contig of 13012 bp in length
 13013 13112: gap of 100 bp
 13113 24677: contig of 11565 bp in length
 24678 24777: gap of 100 bp
 24778 34640: contig of 9863 bp in length
 34641 34740: gap of 100 bp
 34741 43196: contig of 8456 bp in length
 43197 43296: gap of 100 bp
 43297 50126: contig of 6830 bp in length
 50127 50226: gap of 100 bp
 50227 57725: contig of 7499 bp in length
 57726 57826: gap of 100 bp
 57826 64612: contig of 6287 bp in length
 64613 64712: gap of 100 bp
 64713 70999: contig of 6287 bp in length
 71000 71099: gap of 100 bp
 71099 76188: contig of 5089 bp in length
 76189 76288: gap of 100 bp
 76289 83680: contig of 7292 bp in length
 83681 83680: gap of 100 bp
 83681 90511: contig of 6831 bp in length
 90512 90611: gap of 100 bp
 90612 95493: contig of 4882 bp in length
 95494 95593: gap of 100 bp
 95594 102262: contig of 6669 bp in length
 102263 102362: gap of 100 bp
 102363 105866: contig of 3504 bp in length
 105867 105966: gap of 100 bp
 105967 111142: contig of 5176 bp in length
 111143 11142: gap of 100 bp
 11143 115923: contig of 4681 bp in length
 115924 116023: gap of 100 bp
 116024 121747: contig of 5724 bp in length
 121748 125332: contig of 3685 bp in length
 125333 125633: gap of 100 bp
 125633 127927: contig of 2295 bp in length
 127928 128027: gap of 100 bp
 128028 131184: contig of 3157 bp in length
 131185 131284: gap of 100 bp
 131285 132725: contig of 1441 bp in length
 132726 132825: gap of 100 bp
 132826 135566: contig of 2741 bp in length
 135567 135666: gap of 100 bp
 135667 138000: contig of 2334 bp in length
 138001 138100: gap of 100 bp
 138101 141194: contig of 3094 bp in length
 141195 141294: gap of 100 bp
 141295 143619: contig of 2325 bp in length
 143620 143719: gap of 100 bp
 143720 146448: contig of 2729 bp in length
 146449 146548: gap of 100 bp
 146549 149531: contig of 2983 bp in length
 149532 149631: gap of 100 bp
 149632 152157: contig of 2526 bp in length
 152158 152257: gap of 100 bp
 152258 155623: contig of 3366 bp in length
 155624 155723: gap of 100 bp
 155724 158334: contig of 2611 bp in length
 158335 158434: gap of 100 bp
 158435 160509: contig of 2075 bp in length
 160510 160609: gap of 100 bp
 160610 162343: contig of 1734 bp in length

* 162344 162443: gap of 100 bp
* 162444 contig of 2366 bp in length
* 164810 164809: gap of 100 bp
* 164910 167036: contig of 2127 bp in length
* 167037 167136: gap of 100 bp
* 167137 169169: contig of 2033 bp in length
* 169170 169269: gap of 100 bp
* 169270 171052: contig of 1783 bp in length
* 171053 171152: gap of 100 bp
* 171153 172933: contig of 1781 bp in length
* 172934 173033: gap of 100 bp
* 173034 174180: contig of 1147 bp in length
* 174181 174280: gap of 100 bp
* 174281 175729: contig of 1449 bp in length
* 175730 175829: gap of 100 bp
* 175830 176906: contig of 1076 bp in length
* 176906 177005: gap of 100 bp
* 177006 178431: contig of 1426 bp in length.

FEATURES

source
1. 178431
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q24"
/clones="RP11-158J21"
1. 13012
/note="assembly_fragment"
13113. 24677
/note="assembly_fragment"
24778. 34640
/note="assembly_fragment"
34741. 43196
/note="assembly_fragment"
43297. 50126
/note="assembly_fragment"
50227. 57725
/note="assembly_fragment"
57826. 64612

Query Match

Best Local Similarity 99.7%; Pred. No. 4.1e-138;

Matches 382; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 49 GGGCGGAGTGAAGCTGGAGTTCGGGGTGGGGGGAGGCGACTGCTCCGTGGTGTCTGA 108
DB 111412 GGGCGGAGTGAAGCTGGAGTTCGGGGTGGGGGGAGGCGACTGCTCCGTGGTGTCTGA 111471
QY 109 GCGCGGCGAGAGCGGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGGAGGTCCAGT 168
DB 111472 GCGCGGCGAGAGCGGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGGAGGTCCAGT 111531
QY 169 GGGGTGCTTAGG-CCCAAGCCCCCACCAGGCTCCAAAGCTCCAGGCGCTCCCGCAG 227
DB 111532 GGGGTGCTTAGGCGCCCAAGCCCCCACCAGGCTCCAAAGCTCCAGGCGCTCCCGCAG 111591
QY 228 GCACCGTGTCTGCGCCCTCTCTTGGTTCAGAAAGTCCGCCCTGGGGGCGAGTTCGTCCCA 287
DB 111592 GCACCGTGTCTGCGCCCTCTCTTGGTTCAGAAAGTCCGCCCTGGGGGCGAGTTCGTCCCA 111651
QY 288 AAGGTTTCTTCGAAAGAATCTGAGAGGGGCGAGTTCCTTGACCGAGGGAATCTCTGTG 347
DB 111652 AAGGTTTCTTCGAAAGAATCTGAGAGGGGCGAGTTCCTTGACCGAGGGAATCTCTGTG 111711
QY 348 TAGCTTGAAGCGCGCCAGCCCGAGAGATGCTTCCTTCAATAGATGTTTCCCTCGGC 407
DB 111712 TAGCTTGAAGCGCGCCAGCCCGAGAGATGCTTCCTTCAATAGATGTTTCCCTCGGC 111771
QY 408 TTCTCTGCTTATCTACTGGG 430
DB 111772 TTCTCTGCTTATCTACTGGG 111794

RESULT 14
AC024604/c
LOCUS
DEFINITION
Homo sapiens chromosome 10 clone RP11-331G19, WORKING
DRAFT SEQUENCE, 22 unordered pieces.
AC024604
AC024604.3 GI:8389428
HTG: HTGS PHASE1; HTGS_DRAFT; HTGS_CANCELLED.
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Smith, D.R.
1 (bases 1 to 172546)
Genome Therapeutics Corporation Sequencing Center: Human Genome
Sequence Data
Unpublished
2 (bases 1 to 172546)
Smith, D.R.
Direct Submission
Submitted (01-MAR-2000) Genome Therapeutics Corporation, 100 Beaver
Street, Walcham, MA 02453, USA
On Jun 9, 2000 this sequence version replaced gi:7549605.

Genome Center
Center: Genome Therapeutics Corporation
Center code: GTC
Web site: http://www.genomecorp.com/
Contact: gtc-seqcenter@genomecorp.com

Project Information
Center project name: Hg202

Summary Statistics
Sequencing vector: N/A
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 990315
Consensus quality: 152033 bases at least Q40
Consensus quality: 163535 bases at least Q30
Consensus quality: 165256 bases at least Q20
Insert size: 170446; sum-of-contigs
Quality coverage: 4.3x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 22 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 1339: contig of 1339 bp in length
* 1340 1439: gap of unknown length
* 1440 2789: contig of 1350 bp in length
* 2790 2889: gap of unknown length
* 2890 4562: contig of 1673 bp in length
* 4563 4662: gap of unknown length
* 4663 6344: contig of 1682 bp in length
* 6345 6444: gap of unknown length
* 6445 9221: contig of 2777 bp in length
* 9222 9321: gap of unknown length
* 9322 11997: contig of 2676 bp in length
* 11998 12097: gap of unknown length
* 12098 14867: contig of 2770 bp in length
* 14868 14967: gap of unknown length
* 14968 19919: contig of 4952 bp in length
* 19920 20019: gap of unknown length
* 20020 24125: contig of 4106 bp in length
* 24126 24225: gap of unknown length
* 24226 29444: contig of 5219 bp in length
* 29445 29544: gap of unknown length
* 29545 34568: contig of 5024 bp in length
* 34569 39414: contig of 4746 bp in length
* 39415 45885: gap of unknown length
* 45886 45885: contig of 6371 bp in length

| | | | | |
|----|------|--|--|------|
| Db | 531 | | GAGATGATATATTCTACAGAAAGGTCTCAAAAGCCGAAGAGGCAGCCCAAGAAACGCG | 590 |
| Qy | 961 | | TCTGACTACCTTGCCATCCCATCTGAGAAACAGGAGAACTCTGCGGTACCAAGTGAGGAA | 1020 |
| Db | 591 | | TCTGACTACCTTGCCATCCCATCTGAGAAACAGGAGAACTCTGCGGTACCAAGTGAGGAA | 650 |
| Qy | 1021 | | TAGAAACAGGAGCAGTGTGACATGAGGTGGCCTGAAACACCTGAGGGACTGGACATCCCATG | 1080 |
| Db | 651 | | TAGAAACAGGAGCAGTGTGACATGAGGTGGCCTGAAACACCTGAGGGACTGGACATCCCATG | 710 |
| Qy | 1081 | | TTCAGCAATGTCATGCGATCAGGAGG | 1108 |
| Db | 711 | | TTCAGCAATGTCATGCGATCAGGAGG | 738 |

Search completed: April 7, 2005, 07:44:34
 Job time : 5465 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2005, 04:40:54 ; Search time 701 Seconds
(without alignments)
10648.779 Million cell updates/sec

Title: US-09-977-579-4
Perfect score: 1261
Sequence: 1 cctcccttcgagctgagc.....tgscagaactgagaagccgg 1261

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 10

Total number of hits satisfying chosen parameters: 2478043

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001s:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 1261 | 100.0 | 1261 | 3 | AAC67837 Human bet |
| 2 | 1261 | 100.0 | 1261 | 10 | ACF57870 Human SCN |
| 3 | 1210 | 96.0 | 1261 | 10 | ADB78651 Human ion |
| 4 | 1104 | 87.5 | 4052 | 6 | ABA93727 Human sig |
| 5 | 860 | 68.2 | 1510 | 4 | Aaf84146 Human nov |
| 6 | 798 | 63.3 | 978 | 4 | Aah98320 Human EST |
| 7 | 798 | 63.3 | 978 | 13 | ADs11487 |
| 8 | 787 | 62.4 | 1045 | 4 | AAK52345 Human the |
| 9 | 760 | 60.3 | 953 | 13 | ADs10151 |
| 10 | 396 | 31.4 | 1195 | 5 | AA86764 DNA encod |
| 11 | 391 | 31.0 | 3531 | 5 | AA86763 DNA encod |
| 12 | 324 | 25.7 | 471 | 2 | AAV86895 EST clone |
| 13 | 218 | 17.3 | 621 | 5 | AA86762 DNA encod |
| 14 | 216 | 17.1 | 4625 | 4 | AA104971 Human rep |
| 15 | 216 | 17.1 | 4625 | 4 | ABL97864 Human tes |
| 16 | 91 | 7.2 | 407 | 4 | ABA08942 Human vol |
| 17 | 91 | 7.2 | 407 | 4 | AAK53329 Human pol |
| 18 | 74 | 5.9 | 509 | 4 | AAI93385 Human pol |
| 19 | 44 | 3.5 | 53 | 12 | ADQ07690 PCR prime |
| 20 | 44 | 3.5 | 60 | 12 | ADQ07691 PCR prime |

| | | | | | |
|----|----|-----|-------|----|--------------------|
| 21 | 38 | 3.0 | 645 | 5 | AAC90601 Rat sodiu |
| 22 | 38 | 3.0 | 2220 | 3 | AAC67836 Rat beta3 |
| 23 | 38 | 3.0 | 2632 | 5 | AAC90602 Rat sodiu |
| 24 | 38 | 3.0 | 3108 | 5 | AAC90600 Rat sodiu |
| 25 | 33 | 2.6 | 574 | 4 | AA101292 Human rep |
| 26 | 33 | 2.6 | 574 | 4 | ABL96750 Human tes |
| 27 | 28 | 2.2 | 36 | 4 | Aaf84148 Human nov |
| 28 | 24 | 1.9 | 24 | 10 | ADG19077 Voltage g |
| 29 | 23 | 1.8 | 493 | 13 | ADQ55080 Novel can |
| 30 | 22 | 1.7 | 22 | 10 | ADG19076 Voltage g |
| 31 | 22 | 1.7 | 462 | 6 | ABS05386 Human gen |
| 32 | 21 | 1.7 | 1990 | 2 | AAQ25816 Villin ge |
| 33 | 21 | 1.7 | 2152 | 13 | ADR89479 Apoptosis |
| 34 | 35 | 21 | 2352 | 4 | AAK82355 Human imm |
| 35 | 21 | 1.7 | 11122 | 6 | ABX04611 Human end |
| 36 | 20 | 1.6 | 420 | 3 | AAA49871 Mycobacte |
| 37 | 20 | 1.6 | 627 | 13 | ADQ56620 Novel can |
| 38 | 20 | 1.6 | 628 | 5 | AAH94296 Human foe |
| 39 | 20 | 1.6 | 889 | 12 | ADJ42665 Plant CDN |
| 40 | 20 | 1.6 | 918 | 13 | ADS49483 Bacterial |
| 41 | 20 | 1.6 | 2514 | 8 | ACA38647 Prokaryot |
| 42 | 20 | 1.6 | 2517 | 4 | AAH51950 Mycobacte |
| 43 | 20 | 1.6 | 2517 | 8 | ACA37596 Prokaryot |
| 44 | 20 | 1.6 | 2517 | 8 | ACA40274 Prokaryot |
| 45 | 20 | 1.6 | 2547 | 8 | ACA38177 Prokaryot |

ALIGNMENTS

RESULT 1
AAC67837
ID AAC67837 standard; cDNA; 1261 BP.

AC AAC67837;

DT 15-FEB-2001 (first entry)

DE Human beta3 cDNA.

XX Human; beta sub-unit; beta3; analgesic; anticonvulsant;
KW cerebroprotective; vasotropic; cardiant; nootropic; cytostatic;
KW dermatological; gene therapy; voltage-gated sodium channel; pain;
KW epilepsy; stroke; ischaemia; heart disease; Jacobsen Syndrome;
KW familial nonchromaffin paraganglioma; phenylketonuria;
KW Charcot Marie Tooth disease; ss.

OS Homo sapiens.

PN WO200063367-A1.

PD 26-OCT-2000.

PF 24-FEB-2000; 2000WO-EP001783.

PR 15-APR-1999; 99US-0129473P.

PA (WARN) WARNER LAMBERT CO.

PI (UYCA-) UNIV CAMBRIDGE TECH SERVICES LTD.

XX Cox P, Dixon A, Jackson A, Morgan K;

DR WPI; 2000-665241/64.

DR P-PSDB; AAB36002.

XX Novel nucleic acids encoding a beta-3 subunit from a voltage-gated sodium channel, and their corresponding polypeptides, useful for detecting and treating sodium channel-associated conditions, e.g. pain, epilepsy and stroke.

PS Claim 10; Page 70-71; 88pp; English.

XX The present sequence is given in the claims of a specification relating

| Query Match | 100.0% | Score 1261 | DB 10 | Length 1261 |
|-----------------------|----------------|--|----------|-------------|
| Best Local Similarity | 100.0% | Pred. No. 0 | | |
| Matches 1261 | Conservative 0 | Mismatches 0 | Indels 0 | Gaps 0 |
| Qy | 1 | CCCTCCCTTCGAGCTGAGCTTACCTCTGGGCGCAAAACGAGCGAGCGCAGGGCGCGAGTGG | 60 | |
| Db | 1 | CCCTCCCTTCGAGCTGAGCTTACCTCTGGGCGCAAAACGAGCGAGCGCAGGGCGCGAGTGG | 60 | |
| Qy | 61 | AAGCTGAGATTCCGGGTGGCGGGAGGCGCACTGTCTCGTGGTGTGAGCGCCGCGGAGA | 120 | |
| Db | 61 | AAGCTGAGATTCCGGGTGGCGGGAGGCGCACTGTCTCGTGGTGTGAGCGCCGCGGAGA | 120 | |
| Qy | 121 | GGCGGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGGAGGTCAGTGGGGTGCCTTAG | 180 | |
| Db | 121 | GGCGGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGGAGGTCAGTGGGGTGCCTTAG | 180 | |
| Qy | 181 | GGCCAAAGCCCCACCCGGCTCCAAAGCTCCAGGGCTCCCGACGACACGGTGTCTG | 240 | |
| Db | 181 | GGCCAAAGCCCCACCCGGCTCCAAAGCTCCCGAGGCTCCCGACGACACGGTGTCTG | 240 | |
| Qy | 241 | GCCCTTCCTTCGCTCAGAAAGTCCGCCCTGGGGGCAAGTTCGTCTCCAAAGGTTTCCTCG | 300 | |
| Db | 241 | GCCCTTCCTTCGCTCAGAAAGTCCGCCCTGGGGGCAAGTTCGTCTCCAAAGGTTTCCTCG | 300 | |
| Qy | 301 | AAAGAACTCGAGGGCGCAGTCTTGACCGAGGGAATCTCTGTGTGAGCTTTGGAAGC | 360 | |
| Db | 301 | AAAGAACTCGAGGGCGCAGTCTTGACCGAGGGAATCTCTGTGTGAGCTTTGGAAGC | 360 | |
| Qy | 361 | CGCCAGCCCCAGAAAGATGCCTTCCTCAATAGATGTGTTTCCCTGGCTTCTCTCGTGCTT | 420 | |
| Db | 361 | CGCCAGCCCCAGAAAGATGCCTTCCTCAATAGATGTGTTTCCCTGGCTTCTCTCGTGCTT | 420 | |
| Qy | 421 | ATCTACTGGGTGAGTGCTGCTTCCCTGTGTGTGGAAGTCCCTCGGAGACGGAGCC | 480 | |
| Db | 421 | ATCTACTGGGTGAGTGCTGCTTCCCTGTGTGTGGAAGTCCCTCGGAGACGGAGCC | 480 | |
| Qy | 481 | GTGAGGGCAACCCCATGAAGCTCGCTGCACTCTCTGCAATGAAGAGAGAGAGGTGGAG | 540 | |
| Db | 481 | GTGAGGGCAACCCCATGAAGCTCGCTGCACTCTCTGCAATGAAGAGAGAGAGGTGGAG | 540 | |
| Qy | 541 | GCCACACGGTGGTGGAAATGGTTCTACAGGCCGAGGGCGGTAAGATTCTTATTATAC | 600 | |
| Db | 541 | GCCACACGGTGGTGGAAATGGTTCTACAGGCCGAGGGCGGTAAGATTCTTATTATAC | 600 | |
| Qy | 601 | GAGTATCGGAATGGCCACAGAGGTGGAGACCCCTTTACAGGGCGCTCGCAGTGGAAAT | 660 | |
| Db | 601 | GAGTATCGGAATGGCCACAGAGGTGGAGACCCCTTTACAGGGCGCTCGCAGTGGAAAT | 660 | |
| Qy | 661 | GGCAGCAAGGACCTGCGAGGAGGTGCCATCATCTGTGCTCAAAGTCACTCTGAAACGACTCT | 720 | |
| Db | 661 | GGCAGCAAGGACCTGCGAGGAGGTGCCATCATCTGTGCTCAAAGTCACTCTGAAACGACTCT | 720 | |
| Qy | 721 | GGCCTCTACACTGCAATGTGTCGGGAGTTTCGATTTGAGGCGCATCGGCCCTTTGTG | 780 | |
| Db | 721 | GGCCTCTACACTGCAATGTGTCGGGAGTTTCGATTTGAGGCGCATCGGCCCTTTGTG | 780 | |
| Qy | 781 | AAGACGACGGCGCTGATCCCTTAAGAGTCAACGAGGAGGCTGGAGAGGATTCACCTCT | 840 | |

PT Identifying predisposition to an ion channel dysfunction, such as
PT periodic paralysis, cardiac arrhythmias, migraine, Alzheimer's disease,
PT schizophrenia, anxiety and depression, by detecting encoding-gene
XX mutation events.
PS Claim 6; SEQ ID NO 22; 106bp; English.
XX
CC The invention relates to a novel method for identifying a subject
CC predisposed to a disorder associated with ion channel dysfunction. The
CC method comprises ascertaining if at least one of the genes encoding ion
CC channel subunits (ICs) has undergone a mutation event so that a cDNA
CC derived from the subject has any of 134 nucleotide sequences. The method
CC of the invention has nootropic, neuroprotective, inotropic, antipyretic,
CC antitarrhythmic, antimigraine, antidepressant, antiparkinsonian,
CC neuroleptic, tranquiliser, analgesic, nephroretropic, antidiabetic, and
CC ophthalmological activity. A polynucleotide of the invention acts as an
CC ion channel agonist, or ion channel antagonist. The methods, isolated
CC nucleic acids, polypeptides, antibody, selective agonist, antagonist or
CC modulator of an ion channel, cells and genetically modified non-human
CC animal, are useful for the diagnosis and treatment of epilepsy and/or a
CC disorder associated with ion channel dysfunction, such as hyper- or hypo-
CC kalemic periodic paralysis, myotonias, malignant hyperthermia,
CC myasthenia, cardiac arrhythmias, episodic ataxia, migraine, Alzheimer's
CC disease, Parkinson's disease, schizophrenia, hyperkplexia, anxiety,
CC depression, phobic obsessive symptoms, neuropathic pain, inflammatory
CC pain, chronic/acute pain, Bartter's syndrome, polycystic kidney disease,
CC Dent's disease, hyperinsulinaemic hypoglycaemia of infancy, cystic
CC fibrosis, congenital stationary night blindness and total colour
CC blindness. The present sequence represents a mutant cDNA of the
CC invention. The sequence data for this patent is not represented in the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pat_sequences.
XX
SQ Sequence 1261 BP; 263 A; 365 C; 372 G; 261 T; 0 U; 0 Other;

Query Match 96.0%; Score 1210; DB 10; Length 1261;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1260; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCTCCCTCCAGCTGAGCTTACCTGGCGCAACGAGCGAGGCGCGGCGAGTGG 60
DB 1 CCCTCCCTCCAGCTGAGCTTACCTGGCGCAACGAGCGAGGCGCGGCGAGTGG 60

QY 61 AGCTGGAGTCCGGGTGGGCGGAGCGGCTGCTGGTGGTGGTGGAGCGCGCGGAGA 120
DB 61 AGCTGGAGTCCGGGTGGGCGGAGCGGCTGCTGGTGGTGGTGGAGCGCGCGGAGA 120

QY 121 GCGGCGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGAGGCTCCAGTGGGCTCGCTTAG 180
DB 121 GCGGCGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGAGGCTCCAGTGGGCTCGCTTAG 180

QY 181 GCGGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGAGGCTCCAGTGGGCTCGCTTAG 240
DB 181 GCGGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGAGGCTCCAGTGGGCTCGCTTAG 240

QY 241 GCGGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGAGGCTCCAGTGGGCTCGCTTAG 300
DB 241 GCGGCGGAGCGGCTGATCGGCTCCCTCGAACTGGGAGGCTCCAGTGGGCTCGCTTAG 300

QY 301 AAAGAATCTGAGAGGCGGAGCTTCTTGAACGAGGGAATCTCTGTGTAGCTTGGAGGC 360
DB 301 AAAGAATCTGAGAGGCGGAGCTTCTTGAACGAGGGAATCTCTGTGTAGCTTGGAGGC 360

QY 361 CGCGAGCGGCGGAGGCTGCTGCTTCAATAGATGTTTCCCTGGCTTCTCTGCTGCTT 420
DB 361 CGCGAGCGGCGGAGGCTGCTGCTTCAATAGATGTTTCCCTGGCTTCTCTGCTGCTT 420

QY 421 ATCTACTGGGTGCTGCTGCTTCCCTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
DB 421 ATCTACTGGGTGCTGCTGCTTCCCTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480

QY 481 GTGCGAGGCGCAACCCCATGAAGCTGGCTGCATCTCTCTGCATGAAGAGAGGAGTGGAG 540
DB 481 GTGCGAGGCGCAACCCCATGAAGCTGGCTGCATCTCTCTGCATGAAGAGAGGAGTGGAG 540

DB 481 GTGCGAGGCGCAACCCCATGAAGCTGGCTGCATCTCTCTGCATGAAGAGAGGAGTGGAG 540
QY 541 GCCACACGCGTGGTGGAAATGTTCTTACAGGCGCGGCGGTAAGATTTCTTATTATTAC 600
DB 541 GCCACACGCGTGGTGGAAATGTTCTTACAGGCGCGGCGGTAAGATTTCTTATTATTAC 600
QY 601 GAGTATCGGAATGGCCACCAGGAGGTGGAGAGCCCTTTTCAGGGGCGCTGCAGTGGAAAT 660
DB 601 GAGTATCGGAATGGCCACCAGGAGGTGGAGAGCCCTTTTCAGGGGCGCTGCAGTGGAAAT 660
QY 661 GCGCAGCAAGGACCTGCAGAGAGCTGTCCATCACTGTGTCTCAAGCTCACTCTGAACGACTCT 720
DB 661 GCGCAGCAAGGACCTGCAGAGAGCTGTCCATCACTGTGTCTCAAGCTCACTCTGAACGACTCT 720
QY 721 GGCCTCTACCTGCAATGTGTCCGGAGTGTGAGTTGAGTGGCGCATCGGCCCTTTGTG 780
DB 721 GGCCTCTACCTGCAATGTGTCCGGAGTGTGAGTTGAGTGGCGCATCGGCCCTTTGTG 780
QY 781 AAGACGCGCGGCTGATCCCTCAAGAGTCAAGAGGAGGCTGGAGAGGACTTCACTCTCT 840
DB 781 AAGACGCGCGGCTGATCCCTCAAGAGTCAAGAGGAGGCTGGAGAGGACTTCACTCTCT 840
QY 841 GTGGTCTCAGAAATCATGATGTACATCTCTGTGTCTTCTCACCCTGTGGCTGTCTATC 900
DB 841 GTGGTCTCAGAAATCATGATGTACATCTCTGTGTCTTCTCACCCTGTGGCTGTCTATC 900
QY 901 GAGATGATATATGCTACAGAAAGTCTCAAGAGGCTCAAGAGGAGGAGGCGGCAAGAAACGGG 960
DB 901 GAGATGATATATGCTACAGAAAGTCTCAAGAGGCTCAAGAGGAGGAGGCGGCAAGAAACGGG 960
QY 961 TCTGACTACTTGGCTCCATCCCATCTGAGAGCAAGAGGAGAACTCTGCGGTACCAGTGGAGAA 1020
DB 961 TCTGACTACTTGGCTCCATCCCATCTGAGAGCAAGAGGAGAACTCTGCGGTACCAGTGGAGAA 1020
QY 1021 TAGAACAGGAGCAGTGTGACATGAGGTGGCTCGAACACCTGAGGAGCTGGACATCCCATG 1080
DB 1021 TAGAACAGGAGCAGTGTGACATGAGGTGGCTCGAACACCTGAGGAGCTGGACATCCCATG 1080
QY 1081 TTCAGCAATGTCATGGCATCAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140
DB 1081 TTCAGCAATGTCATGGCATCAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1140
QY 1141 ATCCATGCTTCTGTTTCAATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 1200
DB 1141 ATCCATGCTTCTGTTTCAATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 1200
QY 1201 GACTCCCTTAACCTCCATCAGACTTACCGACCATTAAGACTCTGCGAGAACTGAGAGGCGG 1260
DB 1201 GACTCCCTTAACCTCCATCAGACTTACCGACCATTAAGACTCTGCGAGAACTGAGAGGCGG 1260
QY 1261 G 1261
DB 1261 G 1261

RESULT 4
ABA93727
ID ABA93727 standard; cDNA; 4052 BP.
XX
AC ABA93727;
XX
DT 30-APR-2002 (first entry)
XX
DE Human signal transduction cDNA clone amy2_2f18.
XX
KW Human; foetal brain; foetal kidney; melanoma; testis; amygdala;
KW gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200198454-A2.
XX
PD 27-DEC-2001.

XX 25-APR-2001; 2001WO-IB002050.
 PF
 XX
 XX 25-APR-2000; 2000US-0199380P.
 PR
 XX (GSHU-) GERMAN HUMAN GENOME PROJECT.
 PA
 XX
 XX Wiemann S;
 PI
 XX MPI; 2002-055860/07.
 DR P-PSDB; ABB05689.
 DR
 XX Human cDNA sequences and clones derived from human fetal brain, fetal
 PT kidney, melanoma, testis and amygdala cDNA libraries, useful in genetic
 PT screening and therapy.
 PT
 XX
 XX Claim 1; Page 174-175; 611pp; English.
 PS
 XX The present invention describes assemblages and computer readable media
 CC comprising novel human cDNA sequences and clones derived from human
 CC foetal brain, foetal kidney, melanoma, testis and amygdala cDNA
 CC libraries. ABA93702 to ABA93766 represent human cDNA sequences from the
 CC present invention which encode the proteins given in ABB05662 to
 CC ABB05729. The human cDNA sequences and clones can be used in gene
 CC therapy. The clones may be used in a variety of applications, for example
 CC they may be used in profiling assays, for providing large arrays of human
 CC genetic material for implementing large-scale screening strategies and
 CC for treating diseases via gene therapy procedures
 XX
 SQ Sequence 4052 BP; 994 A; 1033 C; 1036 G; 989 T; 0 U; 0 Other;
 Query Match 87.5%; Score 1104; DB 6; Length 4052;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1254; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 3 CTCCTTTCCGAGCTGAGCTTACCTCGGGCGCAACAGCAGGAGGCGGGCGCGAGTGGAA 62
 DB 431 CTCCTTTCCGAGCTGAGCTTACCTCGGGCGCAACAGCAGGAGGCGGGCGCGAGTGGAA 490
 QY 63 GCTGGAGTTCGGGGTGGGGGAGGAGGAGTGTCCGTGGTGTGCTGAGCGCGGGGAGAGC 122
 DB 491 GCTGGAGTTCGGGGTGGGGGAGGAGGAGTGTCCGTGGTGTGCTGAGCGCGGGGAGAGC 550
 QY 123 GGGCGGCGAGCGGCTGATCGGCTCCCTCGAACTGGGAGGTCAGTGGGGTCCGCTTAGGG 182
 DB 551 GGGCGGCGAGCGGCTGATCAGCTCCCTCGAACTGGGAGGTCAGTGGGGTCCGCTTAGGG 610
 QY 183 CCAAGAGCCGCCACCCGGCTCCAAAGCTCCAGGGCTCCCGCAGGACCGGCTGCTCGGC 242
 DB 611 CCAAGAGCCGCCCGGCTCCAAAGCTCCAGGGCTCCCGCAGGACCGGCTGCTCGGC 670
 QY 243 CTTCTCTCGGTGAGAGTGGCCCGGCGGAGTTCGTCGCAAGGGTTTCCTCGAA 302
 DB 671 CTTCTCTCGGTGAGAGTGGCCCGGCGGAGTTCGTCGCAAGGGTTTCCTCGAA 730
 QY 303 AGAATCTGAGAGGCGGAGTCTCTGACCGAGGGAATCTCTGTGTAGCTTGGAGCCG 362
 DB 731 AGAATCTGAGAGGCGGAGTCTCTGACCGAGGGAATCTCTGTGTAGCTTGGAGCCG 790
 QY 363 CCAGCCCGAAGAGTGCCTGCTTCAATAGATTGTTTCCCTGCTTCTCTCGTGTAT 422
 DB 791 CCAGCCCGAAGAGTGCCTGCTTCAATAGATTGTTTCCCTGCTTCTCTCGTGTAT 850
 QY 423 CTACTGGGTGAGTGTCTGCTTCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 482
 DB 851 CTACTGGGTGAGTGTCTGCTTCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 910
 QY 483 GCAGGGCAACCCCATGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 542
 DB 911 GCAGGGCAACCCCATGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 970
 QY 543 CACACCGGTGGTGAATGGTGTCTACAGGCCGAGGGCGGTAAAGATTCTTTATTACGA 602

DB 971 CACCACGGTGGTGAATGGTTCTACAGGCCGAGGGCGGTAAAGATTTCCTTTATTACGA 1030
 QY 603 GTATCGGAATGGCCACCCAGAGGTGGAGAGCCCTTTTCAGGGGCGCTTCGAGTGAATGG 662
 DB 1031 GTATCGGAATGGCCACCCAGAGGTGGAGAGCCCTTTTCAGGGGCGCTTCGAGTGAATGG 1090
 QY 663 CAGCAAGGACCTGCAGGACCTGTCCATCATCTGTGCTCAAGCTCACTCTGAAACGACTCTGG 722
 DB 1091 CAGCAAGGACCTGCAGGACCTGTCCATCATCTGTGCTCAAGCTCACTCTGAAACGACTCTGG 1150
 QY 723 CTTCTACACCTGCAATGTGTCGCGGAGTTTGAAGTTGAGGCGCATCGGCCCTTTGTGAA 782
 DB 1151 CTTCTACACCTGCAATGTGTCGCGGAGTTTGAAGTTTGAAGGCGCATCGGCCCTTTGTGAA 1210
 QY 783 GAGCAGCGGCTGATCCCTTAAGGTCACCGAGGAGGCTGGAGAGGACTTCACCTCTGT 842
 DB 1211 GAGCAGCGGCTGATCCCTTAAGGTCACCGAGGAGGCTGGAGAGGACTTCACCTCTGT 1270
 QY 843 GGTCTCAGAAATCATGATGTACATCTCTTGGTCTTCTCACCCTGTGGCTGTCTCATCGA 902
 DB 1271 GGTCTCAGAAATCATGATGTACATCTCTTGGTCTTCTCACCCTGTGGCTGTCTCATCGA 1330
 QY 903 GATGATATATTGTCTACAGAAAGGTCTCAAAAGCCGAAAGAGGCGAGCCCAAGAAAACGGCTC 962
 DB 1331 GATGATATATTGTCTACAGAAAGGTCTCAAAAGCCGAAAGAGGCGAGCCCAAGAAAACGGCTC 1390
 QY 963 TGACTACCTTGGCCATCCCTTGAAGTCTGAGAACAGGAGAACTTGCCTGTACAGTGGAGGAATA 1022
 DB 1391 TGACTACCTTGGCCATCCCTTGAAGTCTGAGAACAGGAGAACTTGCCTGTACAGTGGAGGAATA 1450
 QY 1023 GAACAGGAGCAGTGTGATGATGAGTGGCTTGAACACCTGAGGGAAGTGGACATCCCATGTT 1082
 DB 1451 GAACAGGAGCAGTGTGATGATGAGTGGCTTGAACACCTGAGGGAAGTGGACATCCCATGTT 1510
 QY 1083 CAGCAATGTCAATGGCATCAGAGGGCGCCCAAGGCGCCCATGCTTCCCTTCATGAT 1142
 DB 1511 CAGCAATGTCAATGGCATCAGAGGGCGCCCAAGGCGCCCATGCTTCCCTTCATGAT 1570
 QY 1143 CCATTGTTCTGTTTCATTTCATTTCATCATATACATCACCCTGCTGTGAGCTTTCACCTCTGA 1202
 DB 1571 CCATTGTTCTGTTTCATTTCATTTCATCATATACATCACCCTGCTGTGAGCTTTCACCTCTGA 1630
 QY 1203 CTCCTTAACCTCCATCAGACCTTACGCAACCATAGACTCTGCCAGAACTGAGAGCC 1259
 DB 1631 CTCCTTAACCTCCATCAGACCTTACGCAACCATAGACTCTGCCAGAACTGAGAGCC 1687
 RESULT 5
 AAF84146
 ID AAF84146 standard; cDNA; 1510 BP.
 XX
 AC AAF84146;
 XX
 DT 07-SEP-2001 (first entry)
 XX
 DE Human novel sodium channel betal-like subunit encoding cDNA.
 XX
 KW Sodium channel; sensory neurone specific channel; betal-like subunit;
 KW SNS; therapeutic; pain; analgesic; ss.
 OS
 XX Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 213..860
 FT /*tag= a
 FT /product= "sodium channel betal-like subunit"
 XX
 XX WO20014293-A2.
 XX
 XX 21-JUN-2001.
 XX
 XX 14-DEC-2000; 2000WO-GB004802.
 XX
 XX

PR 17-DEC-1999; 99CB-00029970.
 XX (GLAX) GLAXO GROUP LTD.
 PA Plumpton M, Powell AJ, Sanseau P;
 PI WPI; 2001-398129/42.
 XX P-PSDB; AAB85206.
 DR Novel sub-unit for voltage-gated sodium channel proteins for producing
 PT agents useful for treating pain.
 XX Claim 4; Page 29-30; 31pp; English.
 PS The invention provides a novel betal-like sub-unit for voltage-gated
 CC sodium ion channel polypeptide, specifically a sensory neurone specific
 CC channel (SNS) subunit. The novel betal-like subunit is useful for
 CC producing a therapeutic agent which is useful treating pain in a patient.
 CC The subunit can be expressed by standard recombinant methodology. The
 CC present sequence represents a human novel sodium channel betal-like
 CC subunit encoding cDNA
 XX
 SQ Sequence 1510 BP; 350 A; 410 C; 414 G; 336 T; 0 U; 0 Other;
 Query Match 68.2%; Score 860; DB 4; Length 1510;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 910; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 349 AGCCTTGAAGCGCCGAGCCAGAGATGCTGCTTCAATAGATGTTTCCCTGGCT 408
 DB 186 AGCCTTGAAGCGCCGAGCCAGAGATGCTGCTTCAATAGATGTTTCCCTGGCT 245
 QY 409 TCTCTGCTTATCTACTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 468
 DB 246 TCTCTGCTTATCTACTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 305
 QY 469 GAGAGCGAGCGCTGAGGCGCAACCCATGAGTGGCTGCTGCTGCTGCTGCTGCTG 528
 DB 306 GAGAGCGAGCGCTGAGGCGCAACCCATGAGTGGCTGCTGCTGCTGCTGCTGCTG 365
 QY 529 GAGAGGCTGAGGCGCCAGCCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 588
 DB 366 GAGAGGCTGAGGCGCCAGCCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 425
 QY 589 TTCCTTTATTAAGATATCGAATATCGAATATCGAATATCGAATATCGAATATCGA 648
 DB 426 TTCCTTTATTAAGATATCGAATATCGAATATCGAATATCGAATATCGAATATCGA 485
 QY 649 CTGCAGTGAATGGCAGCAAGCAGCTGCGAGGAGCTGCTGCTGCTGCTGCTGCTGCT 708
 DB 486 CTGCAGTGAATGGCAGCAAGCAGCTGCGAGGAGCTGCTGCTGCTGCTGCTGCTGCT 545
 QY 709 CTGAACGACTCTGGCTCTTACACCTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 768
 DB 546 CTGAACGACTCTGGCTCTTACACCTGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 605
 QY 769 CGGCCCTTTGAGACGACGCGGCTGATCCCTTAAAGATCAGCAGGAGGCTGAGAG 828
 DB 606 CGGCCCTTTGAGACGACGCGGCTGATCCCTTAAAGATCAGCAGGAGGCTGAGAG 665
 QY 829 GACTTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 888
 DB 666 GACTTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 725
 QY 889 TGGCTGCTCATCGAGATGATATATTGCTACAGAAAGGCTTCAAAAGCCGAGAGGCGCC 948
 DB 726 TGGCTGCTCATCGAGATGATATATTGCTACAGAAAGGCTTCAAAAGCCGAGAGGCGCC 785
 QY 949 CAAGAAACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1008
 DB 786 CAAGAAACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 845
 QY 1009 CCAGTGGAGGAATAGAACAGGAGCAGTGTGACATGAGGTGGCTGAAACCTTGAGGACT 1068

DB 846 CCAGTGGAGGAATAAAGAGGAGCAGTGTGACATGAGGTGGCTGAAACACCTGAGGACT 905
 QY 1069 GGACATCCCATGTTTCCAGCAATGTCAATGGCATCAGAGGGCGCCCAAGGGGCCCATCGC 1128
 DB 906 GGACATCCCATGTTTCCAGCAATGTCAATGGCATCAGAGGGCGCCCAAGGGGCCCATCGC 965
 QY 1129 TTCCCTTCCATGATCCATGTTTCTTCTTCAATTCATTCATCCATCCACCTGCTCTGA 1188
 DB 966 TTCCCTTCCATGATCCATGTTTCTTCTTCAATTCATTCATCCATCCACCTGCTCTGA 1025
 QY 1189 GCTTTTCACTCTGACTTCCCTAACTCCATCAGACCTCTAGGACCATTAAGACTCTGCCAGA 1248
 DB 1026 GCTTTTCACTCTGACTTCCCTAACTCCATCAGACCTCTAGGACCATTAAGACTCTGCCAGA 1085
 QY 1249 ACTGAGAGGCC 1259
 DB 1086 ACTGAGAGGCC 1096
 RESULT 6
 AAH98320
 ID AAH98320 standard; cDNA; 978 BP.
 XX
 AC AAH98320;
 XX
 DT 12-OCT-2001 (first entry)
 XX
 DE Human EST-derived coding sequence SEQ ID NO: 177.
 XX
 KW Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
 KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
 KW diagnostics; forensic test; gene mapping; genetic disorder; biodiversity;
 KW gene therapy; nutrition; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200154477-A2.
 XX
 PD 02-AUG-2001.
 XX
 PF 25-JAN-2001; 2001WO-US002687.
 XX
 PR 25-JAN-2000; 2000US-00491404.
 PR 17-JUL-2000; 2000US-00617746.
 PR 03-AUG-2000; 2000US-00631451.
 PR 15-SEP-2000; 2000US-00663870.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
 PI Cao Y, Drmanac RA, Zhang J, Werhman T;
 XX
 DR WPI; 2001-476164/51.
 DR P-PSDB; AAM23661.
 XX
 PT Isolated polypeptide for treatment of diseases, diagnostics, raising
 PT antibodies and research use.
 XX
 PS Claim 1; Page 299-300; 1275pp; English.
 XX
 CC The present invention provides the protein and coding sequences of novel
 CC proteins from a variety of organisms, including human, dog, cat, horse,
 CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
 CC urchin and tomato. These were derived from expressed sequence tags (ESTs)
 CC from the organism of interest. They can be used in diagnostics,
 CC forensics, gene mapping, identification of mutations, to assess
 CC biodiversity and for nutritional purposes. The present sequence is a cDNA
 CC of the invention
 XX
 SQ Sequence 978 BP; 193 A; 283 C; 293 G; 209 T; 0 U; 0 Other;
 Query Match 63.3%; Score 798; DB 4; Length 978;

| | | Best Local Similarity 100.0%; Pred. No. 0; | | | | Matches 798; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | | | |
|----|-----|--|-----|--|--|--|--|--|--|
| Qy | 88 | GGCGACTGTCGTTGCTGAGCGCGCGGCGGAGAGCGGCGGCGGCTGATCGGCTCC | 147 | | | | | | |
| Db | 7 | GGCGACTGTCGTTGCTGAGCGCGCGGCGGAGAGCGGCGGCTGATCGGCTCC | 66 | | | | | | |
| Qy | 148 | CTCGAACTGGGGAGGTCCAGTGGGGTGCCTTAGGGCCCAAAGCCCCCACCCTCCGCTCCAAA | 207 | | | | | | |
| Db | 67 | CTCGAACTGGGGAGGTCCAGTGGGGTGCCTTAGGGCCCAAAGCCCCCACCCTCCGCTCCAAA | 126 | | | | | | |
| Qy | 208 | AGCTCCAGGCGCTCCCGAGGACCGGTGCTCGGCCCTTCTCTGCGTCAAGAAAGTGCCTC | 267 | | | | | | |
| Db | 127 | AGCTCCAGGCGCTCCCGAGGACCGGTGCTCGGCCCTTCTCTGCGTCAAGAAAGTGCCTC | 186 | | | | | | |
| Qy | 268 | CCTGGGGGAGTTCGTCCTCCAAAGGGTTCTCGAAAGAAATCTGAGAGGGCGGCTCTTG | 327 | | | | | | |
| Db | 187 | CCTGGGGGAGTTCGTCCTCCAAAGGGTTCTCGAAAGAAATCTGAGAGGGCGGCTCTTG | 246 | | | | | | |
| Qy | 328 | ACCGAGGGAATCTCTGCTGAGCCTTGAAGCGCGCCAGCCCGAGAGATGCCTGCTTC | 387 | | | | | | |
| Db | 247 | ACCGAGGGAATCTCTGCTGAGCCTTGAAGCGCGCCAGCCCGAGAGATGCCTGCTTC | 306 | | | | | | |
| Qy | 388 | AATAGATTGTTCCCTGGCTTCTCTGCTTATCTAGGGTCACTGCTGCTCCCT | 447 | | | | | | |
| Db | 307 | AATAGATTGTTCCCTGGCTTCTCTGCTTATCTAGGGTCACTGCTGCTCCCT | 366 | | | | | | |
| Qy | 448 | GTGTGTGGAAGTGCCTCGAGAGCGGCGCTGAGGGCAACCCCATGAAGTGC | 507 | | | | | | |
| Db | 367 | GTGTGTGGAAGTGCCTCGAGAGCGGCGCTGAGGGCAACCCCATGAAGTGC | 426 | | | | | | |
| Qy | 508 | TGCATCTCTGCAATGAAGAGAGAGGAGTGAGGCCACCAACCGTGTGGAATGTTCTAC | 567 | | | | | | |
| Db | 427 | TGCATCTCTGCAATGAAGAGAGAGGAGTGAGGCCACCAACCGTGTGGAATGTTCTAC | 486 | | | | | | |
| Qy | 568 | AGGCCGAGGCGGTAAAGATTTCTTTATACGAGTATCGGAATGGCCACAGGAGGTG | 627 | | | | | | |
| Db | 487 | AGGCCGAGGCGGTAAAGATTTCTTTATACGAGTATCGGAATGGCCACAGGAGGTG | 546 | | | | | | |
| Qy | 628 | GAGAGCCCTTTGAGGGCGCTGAGTGAATGGCAGGAGGACCTGAGGAGTGC | 687 | | | | | | |
| Db | 547 | GAGAGCCCTTTGAGGGCGCTGAGTGAATGGCAGGAGGACCTGAGGAGTGC | 606 | | | | | | |
| Qy | 688 | ATCACTGTGCTCAACGTCACTTGAACGACTCTGGCCCTTACACCTGCAATGTGCCGG | 747 | | | | | | |
| Db | 607 | ATCACTGTGCTCAACGTCACTTGAACGACTCTGGCCCTTACACCTGCAATGTGCCGG | 666 | | | | | | |
| Qy | 748 | GAGTTTGAGTTGAGGCGCATCGGCCCTTTGTGAAGACGACGCGGCTGATCCCTTAAGA | 807 | | | | | | |
| Db | 667 | GAGTTTGAGTTGAGGCGCATCGGCCCTTTGTGAAGACGACGCGGCTGATCCCTTAAGA | 726 | | | | | | |
| Qy | 808 | GTACCCGAGGAGTGGAGAGGACTTACCTCTGTGTCTCAGAAATCATGATGATCATC | 867 | | | | | | |
| Db | 727 | GTACCCGAGGAGTGGAGAGGACTTACCTCTGTGTCTCAGAAATCATGATGATCATC | 786 | | | | | | |
| Qy | 868 | CTTCTGGTCTTCTCACC 885 | | | | | | | |
| Db | 787 | CTTCTGGTCTTCTCACC 804 | | | | | | | |

RESULT 7

| | | |
|----------|----|--|
| ADSL1487 | ID | ADSL1487 standard; DNA; 978 BP. |
| XX | XX | ADSL1487; |
| XX | XX | 16-DEC-2004 (first entry) |
| DT | DT | Human therapeutic contig DNA - SEQ ID 1724. |
| XX | XX | antiinflammatory; neuroprotective; antianaemic; cytostatic; vulnerary; |
| XX | XX | inflammatory; haematopoiesis; immunity; neurodegenerative; stem cell; |
| XX | XX | aplastic anaemia; cancer; wound healing; gene therapy; ds; gene. |

| | | |
|----|----|---|
| XX | OS | Homo sapiens. |
| XX | DN | WO2004080148-A2. |
| XX | PD | 23-SEP-2004. |
| XX | PF | 30-SEP-2003; 2003WO-US030720. |
| XX | PR | 02-OCT-2002; 2002US-0416186P. |
| XX | PA | (NUVE-) NUVELO INC. |
| XX | PI | Tang YT, Asundi V, Ren F, Zhang J, Wehrman T, Wang Z, Ma Y; |
| PI | PI | Wang D, Chen R, Zhao QA, Wang J, Ghosh M, Xue AJ, Weng G, Zhou P; |
| XX | XX | WPI; 2004-668857/65. |
| DR | DR | P-PSDB; ADS12085. |
| XX | XX | New polynucleotide, useful in preparing a composition for diagnosing or |
| PT | PT | treating inflammatory, neurodegenerative or stem cell disorders, e.g., |
| PT | PT | aplastic anemia or cancer for promoting wound healing. |
| XX | XX | Example 2; SEQ ID NO 1724; 718pp; English. |
| PS | PS | The invention relates to a novel isolated polynucleotide and the encoded |
| XX | XX | polypeptide. The molecules of the invention demonstrate antiinflammatory, |
| CC | CC | neuroprotective, antianaemic, cytostatic and vulnerary activities and may |
| CC | CC | be useful in preparing a composition for diagnosing or treating |
| CC | CC | inflammatory, haematopoietic, immune, neurodegenerative or stem cell |
| CC | CC | disorders, such as aplastic anaemia or cancer, as well as for promoting |
| CC | CC | wound healing. The molecules may also be utilised during gene therapy |
| CC | CC | procedures. The current sequence is that of a human therapeutic contig |
| CC | CC | DNA of the invention. The current sequence is not shown explicitly within |
| CC | CC | the specification but can be accessed from the WIPO web-site. |
| XX | XX | Sequence 978 BP; 193 A; 283 C; 293 G; 209 T; 0 U; 0 Other; |

| | | | | | | | | | |
|----|-----|--|-----|--|--|--|--|--|--|
| | | Query Match 63.3%; Score 798; DB 13; Length 978; | | | | Best Local Similarity 100.0%; Pred. No. 0; | | | |
| | | Matches 798; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | | | | | | | |
| Qy | 88 | GGCGACTGTCGTTGCTGAGCGCGCGGCGGAGAGCGGCGGCTGATCGGCTCC | 147 | | | | | | |
| Db | 7 | GGCGACTGTCGTTGCTGAGCGCGCGGCGGAGAGCGGCGGCTGATCGGCTCC | 66 | | | | | | |
| Qy | 148 | CTCGAACTGGGGAGGTCCAGTGGGGTGCCTTAGGGCCCAAAGCCCCCACCCTCCAAA | 207 | | | | | | |
| Db | 67 | CTCGAACTGGGGAGGTCCAGTGGGGTGCCTTAGGGCCCAAAGCCCCCACCCTCCAAA | 126 | | | | | | |
| Qy | 208 | AGCTCCAGGCGCTCCCGAGGACCGGTGCTCGGCCCTTCTCTGCGTCAAGAAAGTGCCTC | 267 | | | | | | |
| Db | 127 | AGCTCCAGGCGCTCCCGAGGACCGGTGCTCGGCCCTTCTCTGCGTCAAGAAAGTGCCTC | 186 | | | | | | |
| Qy | 268 | CCTGGGGGAGTTCGTTCCCAAAGGGTTCTCTGAAAGAAATCTGAGAGGGCGGCTCTTG | 327 | | | | | | |
| Db | 187 | CCTGGGGGAGTTCGTTCCCAAAGGGTTCTCTGAAAGAAATCTGAGAGGGCGGCTCTTG | 246 | | | | | | |
| Qy | 328 | ACCGAGGGAATCTCTGCTGAGCCTTGAAGCGCGCCAGCCCGAGAGATGCCTGCTTC | 387 | | | | | | |
| Db | 247 | ACCGAGGGAATCTCTGCTGAGCCTTGAAGCGCGCCAGCCCGAGAGATGCCTGCTTC | 306 | | | | | | |
| Qy | 388 | AATAGATTGTTCCCTGGCTTCTCTGCTTATCTAGGGTCACTGCTGCTCCCT | 447 | | | | | | |
| Db | 307 | AATAGATTGTTCCCTGGCTTCTCTGCTTATCTAGGGTCACTGCTGCTCCCT | 366 | | | | | | |
| Qy | 448 | GTGTGTGGAAGTGCCTCGAGAGCGGCGCTGAGGGCAACCCCATGAAGTGC | 507 | | | | | | |
| Db | 367 | GTGTGTGGAAGTGCCTCGAGAGCGGCGCTGAGGGCAACCCCATGAAGTGC | 426 | | | | | | |
| Qy | 508 | TGCATCTCTGCAATGAAGAGAGAGGAGTGAGGCCACCAACCGTGTGGAATGTTCTAC | 567 | | | | | | |
| Db | 427 | TGCATCTCTGCAATGAAGAGAGAGGAGTGAGGCCACCAACCGTGTGGAATGTTCTAC | 486 | | | | | | |

QY 568 AGCCCGGAGGCGGTAAAGATTTCTTATTACGATATCGGAATGGCCACAGGAGGTG 627
 Db 487 AGCCCGGAGGCGGTAAAGATTTCTTATTACGATATCGGAATGGCCACAGGAGGTG 546
 QY 628 GAGAGCCCTTTTCAGGGGCGCTGCAAGTGAATGGCAGCAAGACCTGAGGACGTGTCC 687
 Db 547 GAGAGCCCTTTTCAGGGGCGCTGCAAGTGAATGGCAGCAAGACCTGAGGACGTGTCC 606
 QY 688 ATCACTGTGCTCAACTCACTCTGAACGACTCTGGCTCTACACCTGCAATGTGTCGGG 747
 Db 607 ATCACTGTGCTCAACTCACTCTGAACGACTCTGGCTCTACACCTGCAATGTGTCGGG 666
 QY 748 GAGTTGAGTTGAGGCGCATCGCCCTTTGTGAAGACGACGCGGTGATCCCTTAAGA 807
 Db 667 GAGTTGAGTTGAGGCGCATCGCCCTTTGTGAAGACGACGCGGTGATCCCTTAAGA 726
 QY 808 GTCAACGAGGAGGCTGAGAGGACTTCACTCTGTGTCTCGAATATCATGATGATCATC 867
 Db 727 GTCAACGAGGAGGCTGAGAGGACTTCACTCTGTGTCTCGAATATCATGATGATCATC 786
 QY 868 CTTCTGGTCTTCTCACC 885
 Db 787 CTTCTGGTCTTCTCACC 804

RESULT 8

AAK52345
 ID AAK52345 standard; cDNA; 1045 BP.
 AC AAK52345;
 XX
 DT 06-NOV-2001 (first entry)
 DE Human polynucleotide SEQ ID NO 890.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200157190-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; -2001WO-US004098.
 XX
 PR 03-FEB-2000; 2000US-00496914.
 PR 27-APR-2000; 2000US-00560875.
 PR 20-JUN-2000; 2000US-00598075.
 PR 19-JUL-2000; 2000US-00620325.
 PR 01-SEP-2000; 2000US-00654936.
 PR 15-SEP-2000; 2000US-00663561.
 PR 20-OCT-2000; 2000US-00693325.
 PR 30-NOV-2000; 2000US-00728422.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
 PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejman T, Goodrich R;
 XX
 XX WPI; 2001-476283/51.
 DR P-PSDB; AAM79212.
 DR
 XX
 PT Nucleic acids encoding polypeptides with cytokine-like activities, useful
 PT in diagnosis and gene therapy.
 XX
 PS Claim 1; Page 2934-2935; 6221pp; English.
 XX
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the

CC encoded polypeptides' (AAM78323-AAM80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation of which may induce
 CC production of cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
 CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
 CC sequence listing were missing at the time of publication
 XX
 SQ Sequence 1045 BP; 222 A; 289 C; 314 G; 220 T; 0 U; 0 Other;
 Query Match 62.4%; Score 787; DB 4; Length 1045;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 837; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 124 GCGCGGAGCGGCTGATCGGCTCCCTCGAATCGGGAGGTCCAGTGGGTGCTTAGGGC 183
 Db 50 GCGCGGAGCGGCTGATCGGCTCCCTCGAATCGGGAGGTCCAGTGGGTGCTTAGGGC 109
 QY 184 CCAAAAGCCCGCCACCGGCTCCAAAAGCTCCAGGCGCTCCCGAGGCACCGGTGCTCGGC 243
 Db 110 CCAAAAGCCCGCCACCGGCTCCAAAAGCTCCAGGCGCTCCCGAGGCACCGGTGCTCGGC 169
 QY 244 CTTCTTCGCTCAGAAAGTCCGCTCCGCGGCGAGTTCGTCCCAAGGGTTTCTCGAAA 303
 Db 170 CTTCTTCGCTCAGAAAGTCCGCTCCGCGGCGAGTTCGTCCCAAGGGTTTCTCGAAA 229
 QY 304 GAATCTGAGAGGCGCGCTCTTGACCGAGGGAATCTCTGTGTAGCTTGGAGCGCGC 363
 Db 230 GAATCTGAGAGGCGCGCTCTTGACCGAGGGAATCTCTGTGTAGCTTGGAGCGCGC 289
 QY 364 CAGCCCCAGAAAGTCCGCTTCAATAGATTTTCCCTGGCTTCTCTCGTCTATC 423
 Db 290 CAGCCCCAGAAAGTCCGCTTCAATAGATTTTCCCTGGCTTCTCTCGTCTATC 349
 QY 424 TACTGGGTGAGGTCTGTCTTCCCTGTGTGGAAGTCCCTCGGAGACGAGGCGCGT 483
 Db 350 TACTGGGTGAGGTCTGTCTTCCCTGTGTGGAAGTCCCTCGGAGACGAGGCGCGT 409
 QY 484 CAGGCGAACCCTCAAGCTGCGCTGCATCTCTCGATGAGAGAGAGAGAGAGAGAGAG 543
 Db 410 CAGGCGAACCCTCAAGCTGCGCTGCATCTCTCGATGAGAGAGAGAGAGAGAGAGAG 469
 QY 544 ACCACGCTGGTGAATGGTTCTACAGGCGCGAGGCGGTAAAGATTTCTTATTACGAG 603
 Db 470 ACCACGCTGGTGAATGGTTCTACAGGCGCGAGGCGGTAAAGATTTCTTATTACGAG 529
 QY 604 TATCGAATGGCCACAGGAGGTGAGAGCGCCCTTTACAGGGCGCGCTGAGTGAATGCG 663
 Db 530 TATCGAATGGCCACAGGAGGTGAGAGCGCCCTTTACAGGGCGCGCTGAGTGAATGCG 589
 QY 664 AGCAAGGACCTGCAGAGGTGTCATCACTGTGTCTCAAGTCACTGACACCTGACCTGCG 723
 Db 590 AGCAAGGACCTGCAGAGGTGTCATCACTGTGTCTCAAGTCACTGACACCTGACCTGCG 649
 QY 724 CTCTACACCTGCAATGTGTCCCGGAGTTTGAAGTGAAGCGCATCGGCCCTTTGTGAAG 783
 Db 650 CTCTACACCTGCAATGTGTCCCGGAGTTTGAAGTGAAGCGCATCGGCCCTTTGTGAAG 709
 QY 784 ACGACGCGGTGATCCCGCTAAGAGTCAACGAGGAGGTGAGAGACTTCACTCTGTG 843
 Db 710 ACGACGCGGTGATCCCGCTAAGAGTCAACGAGGAGGTGAGAGACTTCACTCTGTG 769
 QY 844 GTCTCAGAAATCATGATGATATCTTCTGTCTTCTCAACCTGCTGCTCATCGAG 903
 Db 770 GTCTCAGAAATCATGATGATATCTTCTGTCTTCTCAACCTGCTGCTCATCGAG 829
 QY 904 ATGATATATTTGCTACAGAAAGGTCTCAAAAGCGGAGAGGCGAGCCCAAGAAACGGT 961

830 ATGATATATTCTACAGAAAGGTCTCAAAAGCCGAGAGGAGCCCAAGAAACGGCT 887

Db

RESULT 9

ADSL0151

ID ADS10151 standard; DNA; 953 BP.

XX AC

XX ADS10151;

XX

DT 16-DEC-2004 (first entry)

XX

DE Human therapeutic DNA - SEQ ID 388.

XX

XX antiinflammatory; neuroprotective; antianaemic; cytostatic; vulnery;

KW inflammatory; haematopoiesis; immunity; neurodegenerative; stem cell;

KW aplastic anaemia; cancer; wound healing; gene therapy; ds; gene.

XX

OS Homo sapiens.

XX

PN WO2004080148-A2.

XX

XX 23-SEP-2004.

XX

PF 30-SEP-2003; 2003WO-US030720.

XX

PR 02-OCT-2002; 2002US-0416186P.

XX

XX (NUVE-) NUVELO INC.

XX

PI Tang YT, Asundi V, Ren F, Zhang J, Wehrman T, Wang Z, Ma Y;

PI Wang D, Chen R, Zhao QA, Wang J, Ghosh M, Xue AJ, Weng G, Zhou P;

XX

XX WPI; 2004-668857/65.

DR P-PSDB; ADS10835.

XX

XX New polynucleotide, useful in preparing a composition for diagnosing or

PT treating inflammatory, neurodegenerative or stem cell disorders, e.g.,

PT aplastic anemia or cancer for promoting wound healing.

XX

PS Claim 1; SEQ ID NO 388; 718pp; English.

XX

CC The invention relates to a novel isolated polynucleotide and the encoded

CC polypeptide. The molecules of the invention demonstrate antiinflammatory,

CC neuroprotective, antianaemic, cytostatic and vulnery activities and may

CC be useful in preparing a composition for diagnosing or treating

CC inflammatory, haematopoietic, immune, neurodegenerative or stem cell

CC disorders, such as aplastic anaemia or cancer, as well as for promoting

CC wound healing. The molecules may also be utilised during gene therapy

CC procedures. The current sequence is that of a human therapeutic DNA of

CC the invention. The current sequence is not shown explicitly within the

CC specification but can be accessed from the WIPO web-site.

XX

SQ Sequence 953 BP; 180 A; 279 C; 290 G; 204 T; 0 U; 0 Other;

Query Match 60.3%; Score 760; DB 13; Length 953;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 760; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 88 GCGGACTGTCGTTGCTGAGCGCCGCGGAGAGCGGCGGCGGCGGCTGATCGGCTCC 147

Db 39 GCGGACTGTCGTTGCTGAGCGCCGCGGAGAGCGGCGGCGGCTGATCGGCTCC 98

Qy 148 CTCGAATCGGGAGGTCAGTGGGGTCTTAGGGCCCAAGCCGCCCGGCTCCAAA 207

Db 99 CTGNACTGGGAGGTCAGTGGGGTCTTAGGGCCCAAGCCGCCCGGCTCCAAA 158

Qy 208 AGCTCCAGGCGCTCCCGGAGCGGCTGCTCGGCGGCTCCCGGAGAGTCCGCC 267

Db 159 AGCTCCAGGCGCTCCCGGAGCGGCTGCTCGGCGGCTCCCGGAGAGTCCGCC 218

Qy 268 CTTGGGGGCGAGTTCGTCGTCGCGGAGGTTCTTCGAAAGAAATCTGAGAGGCGGAGTCTTG 327

Db 219 CTTGGGGGCGAGTTCGTCGTCGCGGAGGTTCTTCGAAAGAAATCTGAGAGGCGGAGTCTTG 278

Qy 328 ACCGAGGGAATCTCTCTGTGTAGCTTTGGAAGCGCGCCAGCCCAAGAGATGCTGCTGCCCTTC 387

Db 279 ACCGAGGGAATCTCTCTGTGTAGCTTTGGAAGCGCGCCAGCCCAAGAGATGCTGCTGCCCTTC 338

Qy 388 AATAGATTGTTTCCCTTGGCTTCTCTCGTCTTATCTACTTGGGTGAGTGTGCTGCTCCCT 447

Db 339 AATAGATTGTTTCCCTTGGCTTCTCTCGTCTTATCTACTTGGGTGAGTGTGCTGCTCCCT 398

Qy 448 GTGTGTGTGAAGTGCCTTCGGAGAGGAGCGGCTGAGGGCAACCCCATGAAGCTGGC 507

Db 399 GTGTGTGTGAAGTGCCTTCGGAGAGGAGCGGCTGAGGGCAACCCCATGAAGCTGGC 458

Qy 508 TGCACTCTCTCTGCAATGAAGAGAGAGGAGTGGAGGCCACACCGGTGGTGGAAATGTTCTAC 567

Db 459 TGCACTCTCTCTGCAATGAAGAGAGAGGAGTGGAGGCCACACCGGTGGTGGAAATGTTCTAC 518

Qy 568 AGGCGCGAGGCGGTAAAGATTTCCTTATTTAGAGTATCGGAATGCGCAACCGAGAGGTG 627

Db 519 AGGCGCGAGGCGGTAAAGATTTCCTTATTTAGAGTATCGGAATGCGCAACCGAGAGGTG 578

Qy 628 GAGAGCGCCCTTCAGGGCGGCTGCAAGTGAAGTGGAGCAAGGACCTGCAGGAGCTGCC 687

Db 579 GAGAGCGCCCTTCAGGGCGGCTGCAAGTGAAGTGGAGCAAGGACCTGCAGGAGCTGCC 638

Qy 688 ATCACTGTGTCTCAACGTCACCTCTGAACGACTCTGGCCCTCTACACCTGCAATGTGTCCGG 747

Db 639 ATCACTGTGTCTCAACGTCACCTCTGAACGACTCTGGCCCTCTACACCTGCAATGTGTCCGG 698

Qy 748 GAGTTGAGTTTGAAGCGCATCGGCCCTTTGTGAAGACGACGCGGCTGATCCCCCTAAGA 807

Db 699 GAGTTGAGTTTGAAGCGCATCGGCCCTTTGTGAAGACGACGCGGCTGATCCCCCTAAGA 758

Qy 808 GTCAACGAGGAGCTGAGAGGAGCTTCACTCTGTGGTCT 847

Db 759 GTCAACGAGGAGCTGAGAGGAGCTTCACTCTGTGGTCT 798

RESULT 10

AAS86764

ID AAS86764 standard; cDNA; 1195 BP.

XX

XX AAS86764;

DT 13-FEB-2002 (first entry)

XX

DE DNA encoding novel human diagnostic protein #22568.

XX

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX

OS Homo sapiens.

XX

PN WO200175067-A2.

XX

PD 11-OCT-2001.

XX

PF 30-MAR-2001; 2001WO-US008631.

XX

PR 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Dmanac RT, Liu C, Tang YT;

XX

XX WPI; 2001-639362/73.

DR P-PSDB; ABG22577.

XX

PT New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity.

| | | | | | | | | | | | | | | | | | | | | | |
|----|-----|-----|-----|----|----|-----|-----|-----|-----|----|----|-----|----|-----|-----|-----|-----|-----|-----|-----|-----|
| Db | 199 | ACG | TGT | TG | GA | TTC | TCT | CAC | GCC | CG | AG | GG | CG | TAA | GAT | TTC | TCT | TAT | TAC | GAT | 258 |
| Qy | 607 | CG | GA | T | CG | CC | AC | CA | GG | AG | TG | GA | GG | CC | CTT | CAG | GG | CG | CG | CT | 666 |
| Db | 259 | CG | GA | T | CG | CC | AC | CA | GG | AG | TG | GA | GG | CC | CTT | CAG | GG | CG | CG | CT | 318 |
| Qy | 667 | AAG | AC | CT | GC | AG | CA | GT | GT | CC | AT | CAC | TG | TG | CT | CAC | TG | AA | CG | ACT | 726 |
| Db | 319 | AAG | AC | CT | GC | AG | CA | GT | GT | CC | AT | CAC | TG | TG | CT | CAC | TG | AA | CG | ACT | 378 |
| Qy | 727 | TAC | AC | CT | GC | AA | TG | TG | CC | GG | AG | TT | TG | AG | TT | TG | AG | GG | CG | CA | 786 |
| Db | 379 | TAC | AC | CT | GC | AA | TG | TG | CC | GG | AG | TT | TG | AG | TT | TG | AG | GG | CG | CA | 438 |
| Qy | 787 | AC | CG | GT | GA | T | CC | CC | CT | TA | AG | AG | T | C | | | | | | | 810 |
| Db | 439 | AC | CG | GT | GA | T | CC | CC | CT | TA | AG | AG | T | C | | | | | | | 462 |

Qy

```

727 TACACCTGCAATGTTGTCCTCGGAGTTCAGTTTGAGCGCCATCGGCCCTTTGTGAAGACG 438
379 TACACCTGCAATGTTGTCCTCGGAGTTCAGTTTGAGCGCCATCGGCCCTTTGTGAAGACG 438
787 ACCGGGCTGATCCCCCTAAGAGTC 810
439 ACGGGCTGATCCCCCTAAGAGTC 462

RESULT 13
AAS86762
ID AAS86762 standard; cDNA; 621 BP.
XX
XX AAS86762;
XX
XX
XX 13-FEB-2002 (first entry)
XX
XX DNA encoding novel human diagnostic protein #22566.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

```

XX Drmanac RT, Liu C, Tang Y[†];
PI
XX
XX WPI; 2001-639362/73.
DR P-PSDB; ABG22575.
DR
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX Claim 1; SEQ ID NO 22566; 103pp; English.

genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and to produce other types of data and products dependent on DNA and

CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 621 BP; 156 A; 155 C; 175 G; 135 T; 0 U; 0 Other;

Query Match 17.3%; Score 218; DB 5; Length 621;
 Best Local Similarity 100.0%; Pred. No. 3.1e-95;
 Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 595 ATTACGATATCGGAATGCCACAGGAGGTGGAGAGCCCTTTTCAGGGGCGCTGCAG 654
 DB |||||
 58 ATTACGATATCGGAATGCCACAGGAGGTGGAGAGCCCTTTTCAGGGGCGCTGCAG 117
 QY 655 TCGAATGGCAGCAAGGACCTGCAGACGCTGCTCATCTCAACGTCACCTCTGAAC 714
 DB |||||
 118 TCGAATGGCAGCAAGGACCTGCAGACGCTGCTCATCTCAACGTCACCTCTGAAC 177
 QY 715 GACTCTGGCCTTACACCTGCAATGTGCCGGAGTTTGAGTTTGAGGCGCATCGGCC 774
 DB |||||
 178 GACTCTGGCCTTACACCTGCAATGTGCCGGAGTTTGAGTTTGAGGCGCATCGGCC 237
 QY 775 TTTGTGAGACACGCGCTGATCCCTTAAGATCAC 812
 DB |||||
 238 TTTGTGAGACACGCGCTGATCCCTTAAGATCAC 275

RESULT 14

AA04971

ID AAL04971 standard; DNA; 4625 BP.

XX AC AAL04971;

XX DT 21-NOV-2001 (first entry)

XX DE Human reproductive system related antigen DNA SEQ ID NO: 7659.

XX KW Human; reproductive system related antigen; reproductive system disorder;
 KW cancer; gene therapy; ds.
 XX OS Homo sapiens.

XX PN WO20015320-A2.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001339.

XX PR 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226868P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235835P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.

Search completed: April 7, 2005, 06:13:24
Job time : 708 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2005, 05:56:24 ; Search time 254 Seconds
(without alignments)
8123.405 Million cell updates/sec

Title: US-09-977-579-4

Perfect score: 1261

Sequence: 1 cctcccttcgagctgagc.....tgccagaactgagaagccgg 1261

Scoring table: Oligo_NUC

Gapop 60.0 , Gapext 60.0

Searched: 1202784 seqs, 818138359 residues

Word size : 10

Total number of hits satisfying chosen parameters: 578812

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

- Issued Patents NA:*
- 1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
 - 2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
 - 3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
 - 4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
 - 5: /cgn2_6/ptodata/1/ina/PCRUS COMB.seq.*
 - 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|---------|----|-------------------------|
| C 1 | 24 | 1.9 | 601 | 4 | US-09-949-016-26733, A |
| C 2 | 24 | 1.9 | 601 | 4 | US-09-949-016-194175, A |
| C 3 | 24 | 1.9 | 137949 | 4 | US-09-949-016-12196, A |
| C 4 | 24 | 1.9 | 137956 | 4 | US-09-949-016-17260, A |
| C 5 | 20 | 1.6 | 1491 | 4 | US-09-902-540-6997, Ap |
| C 6 | 20 | 1.6 | 4312 | 4 | US-09-902-540-612, Ap |
| C 7 | 20 | 1.6 | 4403765 | 3 | US-09-103-840A-2 |
| C 8 | 20 | 1.6 | 4411529 | 3 | US-09-103-840A-1 |
| C 9 | 19 | 1.5 | 601 | 4 | US-09-949-016-85015 |
| C 10 | 19 | 1.5 | 601 | 4 | US-09-949-016-85166 |
| C 11 | 19 | 1.5 | 1626 | 1 | US-08-381-433A-3 |
| C 12 | 19 | 1.5 | 1626 | 4 | US-09-799-978-17 |
| C 13 | 19 | 1.5 | 1626 | 4 | US-09-881-401-3 |
| C 14 | 19 | 1.5 | 5895 | 4 | US-09-902-540-792 |
| C 15 | 19 | 1.5 | 109159 | 4 | US-09-949-016-14169 |
| C 16 | 19 | 1.5 | 109159 | 4 | US-09-949-016-14170 |
| C 17 | 18 | 1.4 | 344 | 3 | US-09-157-177-128 |
| C 18 | 18 | 1.4 | 344 | 4 | US-09-541-210-128 |
| C 19 | 18 | 1.4 | 466 | 3 | US-09-280-116-14 |
| C 20 | 18 | 1.4 | 570 | 4 | US-09-621-976-1635 |
| C 21 | 18 | 1.4 | 601 | 4 | US-09-949-016-56070 |
| C 22 | 18 | 1.4 | 601 | 4 | US-09-949-016-122006 |
| C 23 | 18 | 1.4 | 601 | 4 | US-09-949-016-131295 |
| C 24 | 18 | 1.4 | 601 | 4 | US-09-949-016-131296 |
| C 25 | 18 | 1.4 | 601 | 4 | US-09-949-016-131297 |
| C 26 | 18 | 1.4 | 601 | 4 | US-09-949-016-134047 |
| C 27 | 18 | 1.4 | 601 | 4 | US-09-949-016-138018 |

ALIGNMENTS

RESULT 1

US-09-949-016-26733/c
; Sequence 26733, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26733
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-26733

Query Match 1.9%; Score 24; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|------|-----------------------|------|
| Qy | 1156 | CATTTCATTCATCCATCCACC | 1179 |
| Db | 472 | CATTTCATTCATCCATCCACC | 449 |

RESULT 2

US-09-949-016-194175/c
; Sequence 194175, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498

Sequence 200562,
Sequence 200563,
Sequence 3370, Ap
Sequence 44, Appl
Sequence 9940, Ap
Sequence 10137, A
Sequence 27, Appl
Sequence 1, Appl
Sequence 13374, A
Sequence 15592, A
Sequence 1196, Ap
Sequence 5, Appl
Sequence 11893, A
Sequence 17007, A
Sequence 17389, A
Sequence 13375, A
Sequence 12748, A
Sequence 15165, A

US-09-949-016-200562
US-09-949-016-200563
US-09-902-540-3970
US-09-535-008-44
US-09-252-991A-9940
US-09-252-991A-10137
US-09-023-942A-27
US-08-178-257-1
US-09-949-016-13374
US-09-949-016-15592
US-09-902-540-1196
US-09-146-053-5
US-09-949-016-11893
US-09-949-016-17007
US-09-949-016-17389
US-09-949-016-13375
US-09-949-016-12748
US-09-949-016-15165

18 1.4 601 4
18 1.4 601 4
18 1.4 705 4
18 1.4 713 3
18 1.4 933 4
18 1.4 1023 4
18 1.4 3866 4
18 1.4 4506 4
18 1.4 16914 4
18 1.4 16963 4
18 1.4 18324 4
18 1.4 44453 3
18 1.4 45365 4
18 1.4 45456 4
18 1.4 55031 4
18 1.4 70262 4
18 1.4 70263 4
18 1.4 70947 4

```
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17260
; LENGTH: 137956
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(137956)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17260

Query Match          1.9%; Score 24; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1156 CATTTCATTCATCATCATCCACC 1179
Db 472 CATTTCATTCATCATCATCCACC 449

RESULT 3
US-09-949-016-12196
; Sequence 12196, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12196
; LENGTH: 137949
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(137949)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12196

Query Match          1.9%; Score 24; DB 4; Length 137949;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1156 CATTTCATTCATCATCATCCACC 1179
Db 130759 CATTTCATTCATCATCATCCACC 130782

RESULT 4
US-09-949-016-17260
; Sequence 17260, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17260
; LENGTH: 137956
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(137956)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17260

Query Match          1.9%; Score 24; DB 4; Length 137956;
Best Local Similarity 100.0%; Pred. No. 0.13;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1156 CATTTCATTCATCATCATCCACC 1179
Db 130759 CATTTCATTCATCATCATCCACC 130782

RESULT 5
US-09-902-540-6997
; Sequence 6997, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 6997
; LENGTH: 1491
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-6997

Query Match          1.6%; Score 20; DB 4; Length 1491;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 665 GCAAGGACCTGCAGGACGTG 684
Db 1205 GCAAGGACCTGCAGGACGTG 1224

RESULT 6
US-09-902-540-612/c
; Sequence 612, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 612
; LENGTH: 4312
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-612
```

Query Match 1.6%; Score 20; DB 4; Length 4312;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 665 GCAAGGACCTGCAGGACGTG 684
DB 291 GCAAGGACCTGCAGGACGTG 272

RESULT 7
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: TUBERCULOSIS
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match 1.6%; Score 20; DB 3; Length 4403765;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 500 AGCTGGCTGCATCTCTGTC 519
DB 7383 AGCTGGCTGCATCTCTGTC 7364

RESULT 8
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: TUBERCULOSIS
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match 1.6%; Score 20; DB 3; Length 4411529;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 500 AGCTGGCTGCATCTCTGTC 519

DB 7383 AGCTGGCTGCATCTCTGTC 7364

RESULT 9
US-09-949-016-85015
; Sequence 85015, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 85015
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-85015

Query Match 1.5%; Score 19; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 TGTTCATTCATTCATCCAT 1170
DB 465 TGTTCATTCATTCATCCAT 483

RESULT 10
US-09-949-016-85166
; Sequence 85166, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 85166
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-85166

Query Match 1.5%; Score 19; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 TGTTCATTCATTCATCCAT 1170
DB 465 TGTTCATTCATTCATCCAT 483

RESULT 11

US-08-381-433A-3
; Sequence 3, Application US/08381433A
; Patent No. 5786203
; GENERAL INFORMATION:
; APPLICANT: Lovenberg, Timothy W.
; APPLICANT: Oltersdorf, Tilman
; APPLICANT: Liaw, Chen
; APPLICANT: Grigoriadis, Dimitri E.
; APPLICANT: DeSouza, Errol B.
; TITLE OF INVENTION: CORTICOTROPIN RELEASING FACTOR 2
; TITLE OF INVENTION: RECEPTORS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/381,433A
; FILING DATE: 31-JAN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McMaisters, David D.
; REGISTRATION NUMBER: 33,963
; REFERENCE/DOCKET NUMBER: 690068.401C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; TELEX: 3723836 SEEDANDBERRY
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1626 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 216..1449
US-08-381-433A-3
Query Match 1.5%; Score 19; DB 1; Length 1626;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 72 CCGGGGTGGCGGGGAGGC 90
Db 95 CCGGGGTGGCGGGGAGGC 113
RESULT 12
US-09-799-978-17
; Sequence 17, Application US/09799978
; Patent No. 6670140
; GENERAL INFORMATION:
; APPLICANT: The Procter & Gamble Company
; APPLICANT: Isfort, Robert
; APPLICANT: Sheldon, Russell
; TITLE OF INVENTION: Methods for Identifying Compounds for Regulating Muscle Mass on
; TITLE OF INVENTION: Function Using Corticotropin Releasing Factor Receptors
; FILE REFERENCE: 8448
; CURRENT APPLICATION NUMBER: US/09/799,978
; CURRENT FILING DATE: 2001-03-06
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 1626

; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (216) ..(1451)
US-09-799-978-17
Query Match 1.5%; Score 19; DB 4; Length 1626;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 72 CCGGGGTGGCGGGGAGGC 90
Db 95 CCGGGGTGGCGGGGAGGC 113
RESULT 13
US-09-881-401-3
; Sequence 3, Application US/09881401
; Patent No. 6723841
; GENERAL INFORMATION:
; APPLICANT: Lovenberg, Timothy W.
; APPLICANT: Oltersdorf, Tilman
; APPLICANT: Liaw, Chen
; APPLICANT: Grigoriadis, Dimitri E.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: DeSouza, Errol B.
; TITLE OF INVENTION: CORTICOTROPIN RELEASING FACTOR 2
; TITLE OF INVENTION: RECEPTORS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed Intellectual Property Law Group
; STREET: 701 Fifth Avenue, Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/881,401
; FILING DATE: 13-Jun-2001
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Christiansen, William T.
; REGISTRATION NUMBER: 44,614
; REFERENCE/DOCKET NUMBER: 690068.401C4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1626 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 216..1449
US-09-881-401-3
Query Match 1.5%; Score 19; DB 4; Length 1626;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 72 CCGGGGTGGCGGGGAGGC 90
Db 95 CCGGGGTGGCGGGGAGGC 113

RESULT 14

US-09-902-540-792
; Sequence 792, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 792
; LENGTH: 5895
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-792

Query Match 1.5%; Score 19; DB 4; Length 5895;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 576 GGGCGGTAAAGATTTCCTT 594

Db 5033 GGGCGGTAAAGATTTCCTT 5051

RESULT 15

US-09-949-016-14169
; Sequence 14169, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CU001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14169
; LENGTH: 109159
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14169

Query Match 1.5%; Score 19; DB 4; Length 109159;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 TGTTCATTCATTCATCCAT 1170

Db 50828 TGTTCATTCATTCATCCAT 50846

Search completed: April 7, 2005, 08:54:58
Job time : 266 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2005, 06:02:29 ; Search time 6807 Seconds
(without alignments)
1122.619 Million cell updates/sec

Title: US-09-977-579-4
Perfect score: 1261
Sequence: 1 cctccttcctcgagctgagc.....tgccagaactgagaagccgg 1361

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 5615251 seqs, 3030001701 residues

Word size : 10

Total number of hits satisfying chosen parameters: 3090929

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*

3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*

4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*

5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*

6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*

7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*

8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*

9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*

10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*

11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*

12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*

13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*

14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*

15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*

16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*

17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*

18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*

19: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*

20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*

21: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

22: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|---------------------|
| 1 | 1261 | 100.0 | 1261 | 11 | US-09-977-579-4 |
| 2 | 216 | 17.1 | 4625 | 10 | US-09-764-891-7659 |
| 3 | 91 | 7.2 | 407 | 17 | US-10-276-774-718 |
| 4 | 38 | 3.0 | 645 | 13 | US-10-029-191-21 |
| 5 | 38 | 3.0 | 2220 | 11 | US-09-977-579-3 |
| 6 | 38 | 3.0 | 2632 | 13 | US-10-029-191-22 |
| 7 | 38 | 3.0 | 3108 | 13 | US-10-029-191-1 |
| 8 | 33 | 2.6 | 574 | 10 | US-09-764-891-1293 |
| 9 | 24 | 1.9 | 24 | 16 | US-10-121-759-20 |
| 10 | 22 | 1.7 | 22 | 16 | US-10-121-759-19 |
| 11 | 22 | 1.7 | 462 | 9 | US-09-864-761-15643 |
| | | | | | Sequence 4, Appli |
| | | | | | Sequence 7659, Ap |
| | | | | | Sequence 718, Appl |
| | | | | | Sequence 21, Appl |
| | | | | | Sequence 3, Appli |
| | | | | | Sequence 22, Appl |
| | | | | | Sequence 1, Appli |
| | | | | | Sequence 1293, Ap |
| | | | | | Sequence 20, Appl |
| | | | | | Sequence 19, Appl |
| | | | | | Sequence 15643, A |

| | | | | | | |
|----|----|-----|--------|----|----------------------|-------------------|
| 12 | 21 | 1.7 | 621 | 13 | US-10-027-632-132914 | Sequence 132914, |
| 13 | 21 | 1.7 | 621 | 17 | US-10-027-632-132914 | Sequence 132914, |
| 14 | 20 | 1.6 | 889 | 17 | US-10-260-238-3665 | Sequence 3665, Ap |
| 15 | 20 | 1.6 | 918 | 17 | US-10-369-493-27913 | Sequence 27913, A |
| 16 | 20 | 1.6 | 2514 | 17 | US-10-282-122A-26517 | Sequence 26517, A |
| 17 | 20 | 1.6 | 2517 | 9 | US-09-712-363-4 | Sequence 4, Appli |
| 18 | 20 | 1.6 | 2517 | 17 | US-10-282-122A-25466 | Sequence 25466, A |
| 19 | 20 | 1.6 | 2517 | 17 | US-10-282-122A-28144 | Sequence 28144, A |
| 20 | 20 | 1.6 | 2547 | 17 | US-10-282-122A-26047 | Sequence 26047, A |
| 21 | 20 | 1.6 | 2884 | 14 | US-10-219-219-3 | Sequence 3, Appli |
| 22 | 20 | 1.6 | 58922 | 18 | US-10-322-281-526 | Sequence 526, App |
| 23 | 20 | 1.6 | 200418 | 13 | US-10-087-192-568 | Sequence 568, App |
| 24 | 19 | 1.5 | 201 | 19 | US-10-741-600-68354 | Sequence 68354, A |
| 25 | 19 | 1.5 | 376 | 18 | US-10-425-115-74855 | Sequence 74855, A |
| 26 | 19 | 1.5 | 471 | 9 | US-09-864-761-679 | Sequence 679, App |
| 27 | 19 | 1.5 | 625 | 13 | US-10-027-632-86250 | Sequence 86250, A |
| 28 | 19 | 1.5 | 625 | 13 | US-10-027-632-86251 | Sequence 86251, A |
| 29 | 19 | 1.5 | 625 | 17 | US-10-027-632-86250 | Sequence 86250, A |
| 30 | 19 | 1.5 | 625 | 17 | US-10-027-632-86251 | Sequence 86251, A |
| 31 | 19 | 1.5 | 637 | 13 | US-10-027-632-217363 | Sequence 217363, |
| 32 | 19 | 1.5 | 637 | 17 | US-10-027-632-217363 | Sequence 217363, |
| 33 | 19 | 1.5 | 688 | 18 | US-10-767-701-8737 | Sequence 8737, Ap |
| 34 | 19 | 1.5 | 762 | 13 | US-10-027-632-149779 | Sequence 149779, |
| 35 | 19 | 1.5 | 762 | 17 | US-10-027-632-149779 | Sequence 149779, |
| 36 | 19 | 1.5 | 782 | 13 | US-10-027-632-29933 | Sequence 29933, A |
| 37 | 19 | 1.5 | 782 | 17 | US-10-027-632-29933 | Sequence 29933, A |
| 38 | 19 | 1.5 | 815 | 17 | US-10-425-114-14905 | Sequence 14905, A |
| 39 | 19 | 1.5 | 1626 | 9 | US-09-881-401-3 | Sequence 3, Appli |
| 40 | 19 | 1.5 | 1626 | 10 | US-09-799-978-17 | Sequence 17, Appl |
| 41 | 19 | 1.5 | 1626 | 18 | US-10-649-852-17 | Sequence 17, Appl |
| 42 | 19 | 1.5 | 1626 | 18 | US-10-821-502-3 | Sequence 3, Appli |
| 43 | 19 | 1.5 | 1719 | 17 | US-10-424-599-115665 | Sequence 115665, |
| 44 | 19 | 1.5 | 2024 | 18 | US-10-425-115-77665 | Sequence 77665, A |
| 45 | 19 | 1.5 | 2166 | 17 | US-10-425-114-18441 | Sequence 18441, A |

ALIGNMENTS

RESULT 1

US-09-977-579-4

; Sequence 4, Application US/09977579

; Publication No. US20040248240A1

; GENERAL INFORMATION:

; APPLICANT: Cambridge University Technical Services

; TITLE OF INVENTION: A novel family of beta sub-unit proteins from a voltage gated sod

; TITLE OF INVENTION: channel

; FILE REFERENCE: 674558-2001

; CURRENT APPLICATION NUMBER: US/09/977,579

; CURRENT FILING DATE: 2001-10-15

; PRIOR APPLICATION NUMBER: PCT/EP00/01783

; PRIOR FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: 60,129,473

; PRIOR FILING DATE: 2000-02-24

; NUMBER OF SEQ ID NOS: 47

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 4

; LENGTH: 1261

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-977-579-4

| | | | | |
|-----------------------|-----------------|--|-----------|--------------|
| Query Match | 100.0% | Score 1261; | DB 11; | Length 1261; |
| Best Local Similarity | 100.0%; | Pred. No. 0; | | |
| Matches 1261; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |
| Qy | 1 | CCCTCCCTTCGAGCTGAGCTTACCTCGGCGCAACAGCGAGCGCGGCGGAGTGG | 60 | |
| Db | 1 | CCCTCCCTTCGAGCTGAGCTTACCTCGGCGCAACAGCGAGCGCGGCGGAGTGG | 60 | |
| Qy | 61 | AAGCTGAGTTCGCGGGTGGCGGAGGCGACTGTCCGTGGTCTGAGCCCGCGGAGA | 120 | |

```
Db 61 AAGCTGGAGTTCCGGGGTGGGGGAGGGGAGCTGTCCGTGGTGTGCTGAGGGCGGGCGAGA 120
Qy 121 GGGGGCGGGAGCGGTGATCGGCTCCCTCGAACTGGGAGGTCCAGTGGGGTTCGCTTAG 180
Db 121 GGGGGCGGGAGCGGTGATCGGCTCCCTCGAACTGGGAGGTCCAGTGGGGTTCGCTTAG 180
Qy 181 GGGCCAAAGCCCCACCGGCTCCAAAGCTCCAGGGCTCCCAAGGACACGGTCTCG 240
Db 181 GGGCCAAAGCCCCACCGGCTCCAAAGCTCCAGGGCTCCCAAGGACACGGTCTCG 240
Qy 241 GGCCTTCCTTCGGTCAAGAAAGTCGCCCTCGGGGAGTTCGTCCTCAAGAGGTTTCCTCG 300
Db 241 GGCCTTCCTTCGGTCAAGAAAGTCGCCCTCGGGGAGTTCGTCCTCAAGAGGTTTCCTCG 300
Qy 301 AAAGAAATCTGAGAGGGCGAGTCTTTGACCGGGAATCTCTCTGTGTAGCCTTGAAGC 360
Db 301 AAAGAAATCTGAGAGGGCGAGTCTTTGACCGGGAATCTCTCTGTGTAGCCTTGAAGC 360
Qy 361 CGCCAGCCCCAGAAAGATGCTGCTGCTCAATAGATTGTTTCCCTGCTTCTCTCGTCTT 420
Db 361 CGCCAGCCCCAGAAAGATGCTGCTGCTCAATAGATTGTTTCCCTGCTTCTCTCGTCTT 420
Qy 421 ATCTACTGGTCACTGCTGCTTCCCTGTGTGTGTAAGTGCCTCGGAGCGAGGCC 480
Db 421 ATCTACTGGTCACTGCTGCTTCCCTGTGTGTGTAAGTGCCTCGGAGCGAGGCC 480
Qy 481 GTGAGGGCAACCCCATGAAGTGCCTGCATCTCTGTCATGAAGAGAGAGAGGTGGAG 540
Db 481 GTGAGGGCAACCCCATGAAGTGCCTGCATCTCTGTCATGAAGAGAGAGAGGTGGAG 540
Qy 541 GCCACACCGTGTGAAGTGTCTACAGGCCGAGGGCGGTAAAGTTTCTTATTAC 600
Db 541 GCCACACCGTGTGAAGTGTCTACAGGCCGAGGGCGGTAAAGTTTCTTATTAC 600
Qy 601 GAGTATCGGAATGGCCACAGGAGGTGGAGAGCCCTTTTCAGGGGCGCTGCAGTGGAAT 660
Db 601 GAGTATCGGAATGGCCACAGGAGGTGGAGAGCCCTTTTCAGGGGCGCTGCAGTGGAAT 660
Qy 661 GGCAGCAAGGACTGACAGAGCTGCTCATCTGCTCAAGTCTACTGTAACGACTCT 720
Db 661 GGCAGCAAGGACTGACAGAGCTGCTCATCTGCTCAAGTCTACTGTAACGACTCT 720
Qy 721 GGCCTTACACCTGCAATGTGTCGGGAGTTGAGTTTGGGCGCATCGGCCCTTTTG 780
Db 721 GGCCTTACACCTGCAATGTGTCGGGAGTTGAGTTTGGGCGCATCGGCCCTTTTG 780
Qy 781 AAGACGACGGCTGATCCCTTAAAGTCAAGAGTCAAGGAGGCTGGAGAGACTTCACTCT 840
Db 781 AAGACGACGGCTGATCCCTTAAAGTCAAGAGTCAAGGAGGCTGGAGAGACTTCACTCT 840
Qy 841 GTGGTCTCAGAAATCATGATGATCATCTTCTGGTCTTCTACCTGCTGGCTGCTCATC 900
Db 841 GTGGTCTCAGAAATCATGATGATCATCTTCTGGTCTTCTACCTGCTGGCTGCTCATC 900
Qy 901 GAGATGATATTTGCTACAGAAAGTCTCAAGCGGAGAGGCGACCCCAAGAAACGG 960
Db 901 GAGATGATATTTGCTACAGAAAGTCTCAAGCGGAGAGGCGACCCCAAGAAACGG 960
Qy 961 TCTGACTACTTGCATCCCATCTGAGAAACAGGAACTCTGCGGTACAGTGGAGGAA 1020
Db 961 TCTGACTACTTGCATCCCATCTGAGAAACAGGAACTCTGCGGTACAGTGGAGGAA 1020
Qy 1021 TAGAACAGAGCAGTGTGATGATGAGTGGCTGAAACCTTGAGGAGTCTGGACATCCATG 1080
Db 1021 TAGAACAGAGCAGTGTGATGATGAGTGGCTGAAACCTTGAGGAGTCTGGACATCCATG 1080
Qy 1081 TTCAGCAATGTCAATGGCATCAGAGGCGGCCCAAGGGCCCATCGTTCCTTCATGC 1140
Db 1081 TTCAGCAATGTCAATGGCATCAGAGGCGGCCCAAGGGCCCATCGTTCCTTCATGC 1140
Qy 1141 ATCCATTGTTGTTTCATTTCATTCATCCATACATCCACCTGCTGTGAGCTTTCACCTCT 1200
Db 1141 ATCCATTGTTGTTTCATTTCATTCATCCATACATCCACCTGCTGTGAGCTTTCACCTCT 1200
```

```
Qy 1201 GACTCCCTAACTCCATCAGACCTCTACGCACCATTAAGACTCTGCCAGAACTGAGAAGCCG 1260
Db 1201 GACTCCCTAACTCCATCAGACCTCTACGCACCATTAAGACTCTGCCAGAACTGAGAAGCCG 1260
Qy 1261 G 1261
Db 1261 G 1261
```

RESULT 2

```
US-09-764-891-7659
; Sequence 7659, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7659
; LENGTH: 4625
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-7659
```

Query Match 17.1%; Score 216; DB 10; Length 4625;

Best Local Similarity 100.0%; Pred. No. 5.4e-107; Indels 0; Gaps 0;

Matches 216; Conservative 0; Mismatches 0;

```
Qy 1044 AGGTGGCTTGAACACCTGAGGGACTGGACATCCCATGTTTTCAGCAATGTCAATGGCATCAG 1103
Db 33 AGGTGGCTTGAACACCTGAGGGACTGGACATCCCATGTTTTCAGCAATGTCAATGGCATCAG 92
Qy 1104 GAGGGGGCCCCAAGGGCCCCATCGCTTCCCTTCATGATCCATGTTTCTGTTTCATTTCATT 1163
Db 93 GAGGGGGCCCCAAGGGCCCCATCGCTTCCCTTCATGATCCATGTTTCTGTTTCATTTCATT 152
Qy 1164 CATCCATACATCCACTGCTCTGAGCTTTCACCTCTGACTCCCTAACTCCATCAGACCT 1223
Db 153 CATCCATACATCCACTGCTCTGAGCTTTCACCTCTGACTCCCTAACTCCATCAGACCT 212
Qy 1224 CTACGCACCATTAAGACTCTGCCAGAACTGAGAAGCC 1259
Db 213 CTACGCACCATTAAGACTCTGCCAGAACTGAGAAGCC 248
```

RESULT 3

```
US-10-276-774-718/c
; Sequence 718, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: No. US20040053245A1 Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: US/10/276,774
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 718
; LENGTH: 407
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-276-774-718
```



```
Query Match      7.2%; Score 91; DB 17; Length 407;
Best Local Similarity 99.3%; Pred. No. 1.2e-38;
Matches 141; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 820 GCTGGAGAGACTTCACCTCTGTGGTCTCAGAAATCATGATGACATCCTTCTGGTCTTC 879
Db 230 GCTGGAGAGAGACTTCACCTCTGTGGTCTCAGAAATCATGATGACATCCTTCTGGTCTTC 171

QY 880 CTCACCTGTGGTCTCTCATCGAGATCATATATTGCTACAGAAAGTCTCAAAAGCCGAA 939
Db 170 CTCACCTGTGGTCTCTCATCGAGATCATATATTGCTACAGAAAGTCTCAAAAGCCGAA 111

QY 940 GAGCGAGCCCAAGAAACGGGT 961
Db 110 GAGCGAGCCCAAGAAACGGGT 89

RESULT 4
US-10-029-191-21
; Sequence 21, Application US/10029191
; Publication No. US20020160453A1
; GENERAL INFORMATION:
; APPLICANT: CURTIS, RORY A.J.
; TITLE OF INVENTION: NOVEL GENE ENCODING A SODIUM CHANNEL BETA-3 SUBUNIT
; FILE REFERENCE: 210147.00XX/5U1
; CURRENT APPLICATION NUMBER: US/10/029,191
; CURRENT FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 09/569,978
; PRIOR FILING DATE: 2000-05-12
; PRIOR APPLICATION NUMBER: US 60/134,198
; PRIOR FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 645
; TYPE: DNA
; ORGANISM: Rattus sp.
US-10-029-191-21

Query Match      3.0%; Score 38; DB 13; Length 645;
Best Local Similarity 100.0%; Pred. No. 1e-09;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GTCTGCTTCCCTGTGTGTGGAGTGCCCTCGGAGAC 473
Db 61 GTCTGCTTCCCTGTGTGTGGAGTGCCCTCGGAGAC 98

RESULT 5
US-09-977-579-3
; Sequence 3, Application US/09977579
; Publication No. US20040248240A1
; GENERAL INFORMATION:
; APPLICANT: Cambridge University Technical Services
; TITLE OF INVENTION: A novel family of beta sub-unit proteins from a voltage gated sodium channel
; TITLE OF INVENTION: channel
; FILE REFERENCE: 674558-2001
; CURRENT APPLICATION NUMBER: US/09/977,579
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: PCT/EP00/01783
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60,129,473
; PRIOR FILING DATE: 2000-02-24
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 2220
; TYPE: DNA
; ORGANISM: rat
US-09-977-579-3
```

```
Query Match      3.0%; Score 38; DB 11; Length 2220;
Best Local Similarity 100.0%; Pred. No. 9.6e-10;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GTCTGCTTCCCTGTGTGTGGAGTGCCCTCGGAGAC 473
Db 423 GTCTGCTTCCCTGTGTGTGGAGTGCCCTCGGAGAC 460

RESULT 6
US-10-029-191-22
; Sequence 22, Application US/10029191
; Publication No. US20020160453A1
; GENERAL INFORMATION:
; APPLICANT: CURTIS, RORY A.J.
; TITLE OF INVENTION: NOVEL GENE ENCODING A SODIUM CHANNEL BETA-3 SUBUNIT
; FILE REFERENCE: 210147.00XX/5U1
; CURRENT APPLICATION NUMBER: US/10/029,191
; CURRENT FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 09/569,978
; PRIOR FILING DATE: 2000-05-12
; PRIOR APPLICATION NUMBER: US 60/134,198
; PRIOR FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 2632
; TYPE: DNA
; ORGANISM: Rattus sp.
US-10-029-191-22

Query Match      3.0%; Score 38; DB 13; Length 2632;
Best Local Similarity 100.0%; Pred. No. 9.5e-10;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GTCTGCTTCCCTGTGTGTGGAGTGCCCTCGGAGAC 473
Db 138 GTCTGCTTCCCTGTGTGTGGAGTGCCCTCGGAGAC 175

RESULT 7
US-10-029-191-1
; Sequence 1, Application US/10029191
; Publication No. US20020160453A1
; GENERAL INFORMATION:
; APPLICANT: CURTIS, RORY A.J.
; TITLE OF INVENTION: NOVEL GENE ENCODING A SODIUM CHANNEL BETA-3 SUBUNIT
; FILE REFERENCE: 210147.00XX/5U1
; CURRENT APPLICATION NUMBER: US/10/029,191
; CURRENT FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 09/569,978
; PRIOR FILING DATE: 2000-05-12
; PRIOR APPLICATION NUMBER: US 60/134,198
; PRIOR FILING DATE: 1999-05-14
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 3108
; TYPE: DNA
; ORGANISM: Rattus sp.
US-10-029-191-1

Query Match      3.0%; Score 38; DB 13; Length 3108;
Best Local Similarity 100.0%; Pred. No. 9.4e-10;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 436 GTCTGCTTCCCTGTGTGTGGAGTGCCCTCGGAGAC 473
Db 138 GTCTGCTTCCCTGTGTGTGGAGTGCCCTCGGAGAC 175
```

RESULT 8

US-09-764-891-1293
 ; Sequence 1293, Application US/09764891
 ; Publication No. US20030077808A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PC006
 ; CURRENT APPLICATION NUMBER: US/09/764,891
 ; PRIOR FILING DATE: 2001-01-17
 ; Prior application data removed - consult PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 10231
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1293
 ; LENGTH: 574
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (517)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: SITE
 ; LOCATION: (540)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: SITE
 ; LOCATION: (556)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: SITE
 ; LOCATION: (564)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; NAME/KEY: SITE
 ; LOCATION: (571)
 ; OTHER INFORMATION: n equals a,t,g, or c
 ; US-09-764-891-1293

Query Match 2.6%; Score 33; DB 10; Length 574;
 Best Local Similarity 100.0%; Pred. No. 5.7e-07;
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1227 CGCACCATAAGACTCTGCCAGAACTGAGAAGCC 1259

DB 2 CGCACCATAAGACTCTGCCAGAACTGAGAAGCC 34

RESULT 9

US-10-121-759-20/c
 ; Sequence 20, Application US/10121759
 ; Publication No. US20030194751A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dubin, Adrienne
 ; APPLICANT: Kaftan, Edward
 ; APPLICANT: Chaplan, Sandra
 ; APPLICANT: Brown, Sean
 ; TITLE OF INVENTION: Method For Identifying Modulators Of Ion Channels
 ; FILE REFERENCE: PRI-0007 (ORT 1534)
 ; CURRENT APPLICATION NUMBER: US/10/121,759
 ; CURRENT FILING DATE: 2002-09-26
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 20
 ; LENGTH: 24
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Primer Oligonucleotide
 ; US-10-121-759-20

Query Match 1.9%; Score 24; DB 16; Length 24;
 Best Local Similarity 100.0%; Pred. No. 0.057;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1234 TAAGACTCTGCCAGAACTGAGAAG 1257

|||||

DB

24 TAAGACTCTGCCAGAACTGAGAAG 1

RESULT 10

US-10-121-759-19
 ; Sequence 19, Application US/10121759
 ; Publication No. US20030194751A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dubin, Adrienne
 ; APPLICANT: Kaftan, Edward
 ; APPLICANT: Chaplan, Sandra
 ; APPLICANT: Brown, Sean
 ; TITLE OF INVENTION: Method For Identifying Modulators Of Ion Channels
 ; FILE REFERENCE: PRI-0007 (ORT 1534)
 ; CURRENT APPLICATION NUMBER: US/10/121,759
 ; CURRENT FILING DATE: 2002-09-26
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 19
 ; LENGTH: 22
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Primer Oligonucleotide
 ; US-10-121-759-19

Query Match 1.7%; Score 22; DB 16; Length 22;
 Best Local Similarity 100.0%; Pred. No. 0.71;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 963 TGACTACCTTGCCATCCCATCT 984

DB 1 TGACTACCTTGCCATCCCATCT 22

RESULT 11

US-09-864-761-15643/c
 ; Sequence 15643, Application US/09864761
 ; Patent No. US20020048763A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharron G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; FILE REFERENCE: Aemica-X-1
 ; CURRENT APPLICATION NUMBER: US/09/864,761
 ; CURRENT FILING DATE: 2001-05-23
 ; PRIOR APPLICATION NUMBER: US 60/180,312
 ; PRIOR FILING DATE: 2000-02-04
 ; PRIOR APPLICATION NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 09/632,366
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: GB 24263.6
 ; PRIOR FILING DATE: 2000-10-04
 ; PRIOR APPLICATION NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663
 ; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00662
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: US 60/234,687
 ; PRIOR FILING DATE: 2000-09-21
 ; PRIOR APPLICATION NUMBER: US 09/608,408
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: US 09/774,203
 ; PRIOR FILING DATE: 2001-01-29
 ; NUMBER OF SEQ ID NOS: 49117
 ; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 15643
 ; LENGTH: 462
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO AC011242.1
 ; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
 US-09-864-761-15643

Query Match 1.7%; Score 22; DB 9; Length 462;
 Best Local Similarity 100.0%; Pred. No. 0.58;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1149 TTCTGTTTCATTCATCCAT 1170
 Db 25 TTCTGTTTCATTCATCCAT 4

RESULT 12
 US-10-027-632-132914
 ; Sequence 132914, Application US/10027632
 ; Publication No. US20020198371A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 ; FILE REFERENCE: 108827.129
 ; CURRENT APPLICATION NUMBER: US/10/027,632
 ; CURRENT FILING DATE: 2002-04-30
 ; PRIOR APPLICATION NUMBER: US 60/218,006
 ; PRIOR FILING DATE: 2000-07-12
 ; PRIOR APPLICATION NUMBER: US 60/198,676
 ; PRIOR FILING DATE: 2000-04-20
 ; PRIOR APPLICATION NUMBER: US 60/193,483
 ; PRIOR FILING DATE: 2000-03-29
 ; PRIOR APPLICATION NUMBER: US 60/185,218
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: US 60/167,363
 ; PRIOR FILING DATE: 1999-11-23
 ; PRIOR APPLICATION NUMBER: US 60/156,358
 ; PRIOR FILING DATE: 1999-09-28
 ; PRIOR APPLICATION NUMBER: US 60/146,002
 ; PRIOR FILING DATE: 1999-08-09
 ; NUMBER OF SEQ ID NOS: 325720
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 132914
 ; LENGTH: 621
 ; TYPE: DNA
 ; ORGANISM: Human

Query Match 1.7%; Score 21; DB 13; Length 621;
 Best Local Similarity 100.0%; Pred. No. 2;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1150 TCTGTTTCATTCATCCAT 1170
 Db 194 TCTGTTTCATTCATCCAT 214

RESULT 13
 US-10-027-632-132914
 ; Sequence 132914, Application US/10027632
 ; Publication No. US20030204075A9
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 ; FILE REFERENCE: 108827.129
 ; CURRENT APPLICATION NUMBER: US/10/027,632
 ; CURRENT FILING DATE: 2002-04-30
 ; PRIOR APPLICATION NUMBER: US 60/218,006
 ; PRIOR FILING DATE: 2000-07-12
 ; PRIOR APPLICATION NUMBER: US 60/198,676
 ; PRIOR FILING DATE: 2000-04-20
 ; PRIOR APPLICATION NUMBER: US 60/193,483
 ; PRIOR FILING DATE: 2000-03-29
 ; PRIOR APPLICATION NUMBER: US 60/185,218
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: US 60/167,363
 ; PRIOR FILING DATE: 1999-11-23
 ; PRIOR APPLICATION NUMBER: US 60/156,358
 ; PRIOR FILING DATE: 1999-09-28
 ; PRIOR APPLICATION NUMBER: US 60/146,002
 ; PRIOR FILING DATE: 1999-08-09
 ; NUMBER OF SEQ ID NOS: 325720
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 132914
 ; LENGTH: 621
 ; TYPE: DNA
 ; ORGANISM: Human

Query Match 1.7%; Score 21; DB 17; Length 621;
 Best Local Similarity 100.0%; Pred. No. 2;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1150 TCTGTTTCATTCATCCAT 1170
 Db 194 TCTGTTTCATTCATCCAT 214

RESULT 14
 US-10-260-238-3665
 ; Sequence 3665, Application US/10260238
 ; Publication No. US20040016025A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Budworth, Paul R.
 ; APPLICANT: Moughamer, Todd G.
 ; APPLICANT: Briggs, Steven P.
 ; APPLICANT: Cooper, Bret
 ; APPLICANT: Glazebrook, Jane
 ; APPLICANT: Goff, Stephen A.
 ; APPLICANT: Katagiri, Fumiyaki
 ; APPLICANT: Krepis, Joel
 ; APPLICANT: Provart, Nicholas
 ; APPLICANT: Ricke, Darrell
 ; APPLICANT: Zhu, Tong
 ; TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION
 ; FILE REFERENCE: 60111-NP
 ; CURRENT APPLICATION NUMBER: US/10/260,238
 ; CURRENT FILING DATE: 2002-09-26
 ; PRIOR APPLICATION NUMBER: US 60/325,448
 ; PRIOR FILING DATE: 2001-09-26
 ; PRIOR APPLICATION NUMBER: US 60/325,277
 ; PRIOR FILING DATE: 2001-09-26
 ; PRIOR APPLICATION NUMBER: US 60/370,620
 ; PRIOR FILING DATE: 2002-04-04
 ; NUMBER OF SEQ ID NOS: 6077
 ; SEQ ID NO 3665
 ; LENGTH: 889
 ; TYPE: DNA

; ORGANISM: Triticum aestivum
US-10-260-238-3665

Query Match 1.6%; Score 20; DB 17; Length 889;
Best Local Similarity 100.0%; Pred. No. 6.9; Mismatches 0; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 474 GGAGCCGTGCAGGCAACC 493
|||
Db 318 GGAGCCGTGCAGGCAACC 337

RESULT 15

US-10-369-493-27913/c
; Sequence 27913, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 27913
; LENGTH: 918
; TYPE: DNA
; ORGANISM: Neurospora crassa
US-10-369-493-27913

Query Match 1.6%; Score 20; DB 17; Length 918;
Best Local Similarity 100.0%; Pred. No. 6.9; Mismatches 0; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 110 CGCCGGCGAGCGGCGCG 129
|||
Db 148 CGCCGGCGAGCGGCGCG 129

Search completed: April 7, 2005, 10:53:58
Job time : 6808 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2005, 05:56:04 ; Search time 3942 Seconds
(without alignments)
12176.318 Million cell updates/sec

Title: US-09-977-579-4

Perfect score: 1261

Sequence: 1 cctccctccgagctgagc.....tgccagaactgagaagccgg 1261

Scoring table: OLIGO_NUC

Gapop_60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 10

Total number of hits satisfying chosen parameters: 31203175

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:*

- 1: gb_est1.*
- 2: gb_est2.*
- 3: gb_hic.*
- 4: gb_est3.*
- 5: gb_est4.*
- 6: gb_est5.*
- 7: gb_est6.*
- 8: gb_gss1.*
- 9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 1104 | 87.5 | 4052 | 3 | HSM801563 |
| 2 | 860 | 68.2 | 2555 | 3 | CR609664 |
| 3 | 600 | 47.6 | 1062 | 5 | EX420015 |
| 4 | 597 | 47.3 | 648 | 9 | AY419145 |
| 5 | 499 | 39.6 | 582 | 5 | BP200910 |
| 6 | 443 | 35.1 | 582 | 5 | BP202832 |
| 7 | 437 | 34.7 | 584 | 5 | BP361278 |
| 8 | 384 | 30.5 | 975 | 5 | EX445002 |
| 9 | 381 | 30.2 | 582 | 5 | BP311484 |
| 10 | 348 | 27.6 | 975 | 5 | EX452112 |
| 11 | 288 | 22.8 | 1069 | 5 | BM928131 |
| 12 | 268 | 21.3 | 1078 | 1 | AL534136 |
| 13 | 165 | 13.1 | 443 | 7 | D44825 |
| 14 | 158 | 12.5 | 584 | 9 | AY419146 |
| 15 | 154 | 12.2 | 509 | 5 | EX280396 |
| 16 | 109 | 8.6 | 972 | 5 | EQ066875 |
| 17 | 70 | 5.6 | 926 | 5 | EX411288 |
| 18 | 48 | 3.8 | 646 | 9 | CE606396 |
| 19 | 38 | 3.0 | 471 | 6 | CB732717 |
| 20 | 38 | 3.0 | 472 | 6 | CB730533 |
| 21 | 34 | 2.7 | 176 | 2 | AW832826 |
| 22 | 33 | 2.6 | 631 | 6 | CB577357 |
| 23 | 32 | 2.5 | 439 | 6 | CB750230 |
| 24 | 29 | 2.3 | 636 | 5 | BM933157 |

| | | | | | | |
|----|----|-----|------|---|----------|--------------------|
| 25 | 29 | 2.3 | 648 | 9 | AY419147 | AY419147 Mus muscu |
| 26 | 29 | 2.3 | 652 | 2 | BB652801 | BB652801 BB652801 |
| 27 | 29 | 2.3 | 672 | 2 | BB614118 | BB614118 BB614118 |
| 28 | 29 | 2.3 | 700 | 4 | BT739617 | BT739617 603361873 |
| 29 | 29 | 2.3 | 705 | 6 | CD349958 | CD349958 UI-M-FY0- |
| 30 | 29 | 2.3 | 714 | 7 | CF531573 | CF531573 UI-M-FY0- |
| 31 | 29 | 2.3 | 723 | 6 | CA749311 | CA749311 UI-M-FY0- |
| 32 | 29 | 2.3 | 742 | 6 | CB526211 | CB526211 UI-M-FY0- |
| 33 | 29 | 2.3 | 750 | 6 | CD349206 | CD349206 UI-M-FY0- |
| 34 | 29 | 2.3 | 780 | 5 | BQ770528 | BQ770528 UI-M-FY0- |
| 35 | 29 | 2.3 | 825 | 5 | BQ745919 | BQ745919 UI-M-EX0- |
| 36 | 29 | 2.3 | 845 | 6 | CA327438 | CA327438 UI-M-FY0- |
| 37 | 29 | 2.3 | 846 | 6 | CD355879 | CD355879 UI-M-FY0- |
| 38 | 29 | 2.3 | 927 | 5 | BQ713131 | BQ713131 AGENCOURT |
| 39 | 29 | 2.3 | 950 | 4 | BG294174 | BG294174 602391245 |
| 40 | 29 | 2.3 | 1359 | 3 | AK076466 | AK076466 Mus muscu |
| 41 | 29 | 2.3 | 3549 | 3 | AK049747 | AK049747 Mus muscu |
| 42 | 29 | 2.3 | 4105 | 3 | BC058083 | BC058083 Mus muscu |
| 43 | 29 | 2.3 | 4149 | 3 | AK049286 | AK049286 Mus muscu |
| 44 | 28 | 2.2 | 712 | 6 | CB526257 | CB526257 UI-M-FY0- |
| 45 | 26 | 2.1 | 734 | 7 | CK367344 | CK367344 AGENCOURT |

ALIGNMENTS

RESULT 1
HSM801563
LOCUS Homo sapiens mRNA; cDNA DKFZp761F182 (from clone DKFZp761F182). linear HTC 22-SEP-2004
DEFINITION AL136589
ACCESSION AL136589
VERSION AL136589.1 GI:13276680
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4052)
AUTHORS Ottenwagelder, B., Obermaier, B., Deutschenbaur, S., Schaidp, A., Mewes, H.W., Weil, B., Amid, C., Osanger, A., Fobo, G., Han, M. and Wiemann, S.
CONSRMT The German CDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by Medigenomix (Martinsried/Germany) within the cDNA sequencing consortium of the German Genome Project.
This clone (DKFZp761F182) is available at the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp761F182
Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.
FEATURES
Location/Qualifiers
1..4052
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFZp761F182"
/db_xref="taxon:9606"
/clone="DKFZp761F182"
/tissue type="amygdala"
/clone_lib="761 (synonym: hamy2). Vector pSport1; host DH10B; sites NotI + SalI"
/dev_stage="adult"
/notes="voltage-gated sodium channel beta-3 subunit"
1..4052
/gene="DKFZp761F182"
804..1451
/gene="DKFZp761F182"
/codon_start=1
/product="hypothetical protein"

```
/protein_id="CAB66524.1"
/db_xref="GI:13276881"
/db_xref="GOA:Q9NY72"
/db_xref="UniProt/Swiss-Prot:Q9NY72"
/translation="MPAFNRLPLASLVLIYVSVCFVPCVVEPSETEAVQGNPKMLR
CISMKREVEATTVVWFYRPEGKDFLIYEVRNGHOREVSPFOGRLOWNGSKDLQD
VSTLVNLVTLNDSGLYTCNVSREFEPEARHPFKVTRLLIPLRVTEBAGDFISVSEI
MMIILLVFLTMLLIEIMYCYRKYSKABEAENASDYLAIPTSENKENSAPVVE"
```

ORIGIN

```
Query Match      87.5%; Score 1104; DB 3; Length 4052;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1254; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 CTCCTTCCGAGCTGAGCTTACCTCGGCGCAAAACGAGGAGCGAGGCGCGAGTGGAA 62
DB 431 CTCCTTCCGAGCTGAGCTTACCTCGGCGCAAAACGAGGAGCGAGGCGCGAGTGGAA 490

QY 63 GCTGGAGTTCCGGGTTGGGGGGGAGCGACTGTCTCGTGTGTGAGCGCGCGGAGAGC 122
DB 491 GCTGGAGTTCCGGGTTGGGGGGGAGCGACTGTCTCGTGTGTGAGCGCGCGGAGAGC 550

QY 123 GGGCGCGAGCGGCTGATCGGCTCCCTCGAACTGGGGAGGTCCAGTGGGTCGCTTAGG 182
DB 551 GGGCGCGAGCGGCTGATCAGCTCCCTCGAACTGGGGAGGTCCAGTGGGTCGCTTAGG 610

QY 183 CCAAAAGCCCCACCGGCTCCAAAAGCTCCAGGCGCTCCCGAGCACCGGTGCTCGGC 242
DB 611 CCAAAAGCCCCACCGGCTCCAAAAGCTCCAGGCGCTCCCGAGCACCGGTGCTCGGC 670

QY 243 CTTCTCTTGGTTCAGAAAGTCGCGCCCTGGGGCAGTTGCTCCCAAGGGTTTCCTCGAA 302
DB 671 CTTCTCTTGGTTCAGAAAGTCGCGCCCTGGGGCAGTTGCTCCCAAGGGTTTCCTCGAA 730

QY 303 AGAATCTGAGAGGCGCAGTCCCTTGACCGAGGAATCTCTCTGTGAGCTTGGAGCGC 362
DB 731 AGAATCTGAGAGGCGCAGTCCCTTGACCGAGGAATCTCTCTGTGAGCTTGGAGCGC 790

QY 363 CGAGCCCCAGAAAGATCGCTTCAATAGATTGTTTCCCTCGGCTTCTCTGTGCTTAT 422
DB 791 CGAGCCCCAGAAAGATCGCTTCAATAGATTGTTTCCCTCGGCTTCTCTGTGCTTAT 850

QY 423 CTACTGGGTTCAGTCTGCTTCCCTCTGTCGTGAGAGTGCCTCGAGACGGAGCGCT 482
DB 851 CTACTGGGTTCAGTCTGCTTCCCTCTGTCGTGAGAGTGCCTCGAGACGGAGCGCT 910

QY 483 GCAGGGCAACCCCATGAAGCTGCGCTGCATCTCTGCATGAAGAGAGAGGTGAGGC 542
DB 911 GCAGGGCAACCCCATGAAGCTGCGCTGCATCTCTGCATGAAGAGAGAGGTGAGGC 970

QY 543 CACACGGTGGTGAATGTTTACAGGCCGAGGCGGTAAGATTCTCTTATTTACGA 602
DB 971 CACACGGTGGTGAATGTTTACAGGCCGAGGCGGTAAGATTCTCTTATTTACGA 1030

QY 603 GTATCGGAATGCCACAGAGGTGAGAGCCCTTTTCAGGGCGCTCGAGTGAATCG 662
DB 1031 GTATCGGAATGCCACAGAGGTGAGAGCCCTTTTCAGGGCGCTCGAGTGAATCG 1090

QY 663 CAGCAAGGACCTCAGGAGGTGCTTCACTCACTGCTCAACGTCACTCTGAACGACTCTGG 722
DB 1091 CAGCAAGGACCTCAGGAGGTGCTTCACTCACTGCTCAACGTCACTCTGAACGACTCTGG 1150

QY 723 CTTCTACACTGCAATGTGTCCCGGAGTTTGTAGTTTGAAGCGGCATCGGCCCTTTGTGAA 782
DB 1151 CTTCTACACTGCAATGTGTCCCGGAGTTTGTAGTTTGAAGCGGCATCGGCCCTTTGTGAA 1210

QY 783 GACGACGCGCTGATCCCTTAGAGTCAACGAGGAGGCTGGAGAGACTTACCTCTGT 842
DB 1211 GACGACGCGCTGATCCCTTAGAGTCAACGAGGAGGCTGGAGAGACTTACCTCTGT 1270

QY 843 GGTCTCAGAAATCATGATGTACATCTTCTGTGCTTCTCCTCACCTGTGGTGTCTCATCGA 902
DB 1271 GGTCTCAGAAATCATGATGTACATCTTCTGTGCTTCTCCTCACCTGTGGTGTCTCATCGA 1330
```

```
QY 903 GATGATATATTGCTTACAGAAAGGTCTCAAAAGCCGAGAGGAGGCCAAGAAAAACCGCTC 962
DB 1331 GATGATATATTGCTTACAGAAAGGTCTCAAAAGCCGAGAGGAGGCCAAGAAAAACCGCTC 1390

QY 963 TGACTACCTTGGCATCCCATCTGAGAACAGAGAACTCTGCGGTACCACTGGAGGAATA 1022
DB 1391 TGACTACCTTGGCATCCCATCTGAGAACAGAGAACTCTGCGGTACCACTGGAGGAATA 1450

QY 1023 GAACAGGAGCAGTGTGACATGAGGTGGCTGAACACCTGAGGAGCTTGGACATCCCATGTT 1082
DB 1451 GAACAGGAGCAGTGTGACATGAGGTGGCTGAACACCTGAGGAGCTTGGACATCCCATGTT 1510

QY 1083 CAGCAATGTAATGGCATCAGGAGGCGGCCCAAGGCGCCCATCGCTTCCCTTCATGCAT 1142
DB 1511 CAGCAATGTAATGGCATCAGGAGGCGGCCCAAGGCGCCCATCGCTTCCCTTCATGCAT 1570

QY 1143 CCATTGTTCTGTTCAATTCATTATCATCATACCTCCACCTGCTGAGCTTTCACCTCTGA 1202
DB 1571 CCATTGTTCTGTTCAATTCATTATCATCATACCTCCACCTGCTGAGCTTTCACCTCTGA 1630

QY 1203 CTCCTTAATCCATCAGACCTCTAGCCACCAATAAGACTCTGCCAGAACTTGAGAGCC 1259
DB 1631 CTCCTTAATCCATCAGACCTCTAGCCACCAATAAGACTCTGCCAGAACTTGAGAGCC 1687
```

RESULT 2

```
CR609664
LOCUS      2555 bp      mRNA      linear      HTC 21-JUL-2004
DEFINITION full-length cDNA clone CS0DF023YA09 of Fetal brain of Homo sapiens
(human).
ACCESSION  CR609664
VERSION    CR609664.1 GI:50490471
KEYWORDS   HTC; CNSLT cDNA.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 2555)
AUTHORS   Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE     Full-length cDNA libraries and normalization
JOURNAL    Unpublished
REMARK     Contact : Feng Liang Email : fliang@lifetech.com URL :
            http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
            Faraday Avenue
            Genoscope.
            2 (bases 1 to 2555)
            Direct Submission
            Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT    1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
```

FEATURES

```
source
1..2555
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF023YA09"
/tissue_type="Fetal brain"
/plasmid="pCMVSPORT_6"
```

ORIGIN

```
Query Match      68.2%; Score 860; DB 3; Length 2555;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 910; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 349 AGCCTTGGAGCGCGCAGAGAGTGCCTGCTCAATAGATTGTTCCCTGGCT 408
DB 208 AGCCTTGGAGCGCGCAGAGAGTGCCTGCTCAATAGATTGTTCCCTGGCT 267
```

409 TCTCTGCTTATCTACTGGGTGAGTGTCTGCTTCCCTGTGTGTGGAGTGCCTCG 468
 Db TCTCTGCTTATCTACTGGGTGAGTGTCTGCTTCCCTGTGTGTGGAGTGCCTCG 327
 QY GAGACGAGGCGGTGCGAGGCAACCCCATGAAGCTGCGTGCATCTCTGATGAAGAGA 528
 Db GAGACGAGGCGGTGCGAGGCAACCCCATGAAGCTGCGTGCATCTCTGATGAAGAGA 387
 QY GAGGAGGTGGAGGCGACACGCTGGTGAATGGTTCTACAGGCCGAGGCGGTAAAGAT 588
 Db GAGGAGGTGGAGGCGACACGCTGGTGAATGGTTCTACAGGCCGAGGCGGTAAAGAT 447
 QY TTCTCTTATTTACGAGTATCGGAATGGCCACAGGAGGTGAGAGCCCTTTTCAGGGGCGC 648
 Db TTCTCTTATTTACGAGTATCGGAATGGCCACAGGAGGTGAGAGCCCTTTTCAGGGGCGC 507
 QY CTGACGTGGAATGCGCAGAGGACCTGACAGGACGTGTCCATCACTGTGCTCAACGTCACT 708
 Db CTGACGTGGAATGCGCAGAGGACCTGACAGGACGTGTCCATCACTGTGCTCAACGTCACT 567
 QY CTGAACGACTCTGSCCTCTACACTGCAATGTGTCCCGGAGTTTGAGTTTGAGGCGCAT 768
 Db CTGAACGACTCTGSCCTCTACACTGCAATGTGTCCCGGAGTTTGAGTTTGAGGCGCAT 627
 QY CGGCGCTTTGTGAAGACGACGCGGCTGATCCCCCTAAGAGTCAACGAGGAGGCTGGAGAG 828
 Db CGGCGCTTTGTGAAGACGACGCGGCTGATCCCCCTAAGAGTCAACGAGGAGGCTGGAGAG 687
 QY GACTTCACTCTGTGGTCTCAGAAATCATGATGATGATGATGATGATGATGATGATGATGAT 888
 Db GACTTCACTCTGTGGTCTCAGAAATCATGATGATGATGATGATGATGATGATGATGATGAT 747
 QY TGCTGCTCATCGAGATGATATTTGCTACAGAAAGTCTCAAAAGCCGAGGAGGCGCAT 948
 Db TGCTGCTCATCGAGATGATATTTGCTACAGAAAGTCTCAAAAGCCGAGGAGGCGCAT 807
 QY CAAGAAACCGCTCTGACTACCTTTGCCATCCCTCTCAGAAACGAGGAGTCTTCCGGTA 1008
 Db CAAGAAACCGCTCTGACTACCTTTGCCATCCCTCTCAGAAACGAGGAGTCTTCCGGTA 867
 QY CCAGTGAGGAATAGAAACAGAGCAGTGTGACATGAGTGGCTGCAACCTGAGGAGCT 1068
 Db CCAGTGAGGAATAGAAACAGAGCAGTGTGACATGAGTGGCTGCAACCTGAGGAGCT 927
 QY GGACATCCCATGTTTCAAGTCAATGGCATCAGGAGGCGGCCCAAGGCGCCCATCGC 1128
 Db GGACATCCCATGTTTCAAGTCAATGGCATCAGGAGGCGGCCCAAGGCGCCCATCGC 987
 QY TTCCCTTCATGCTCATTTGTTTCAATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1188
 Db TTCCCTTCATGCTCATTTGTTTCAATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 1047
 QY GCTTTACCTCTGACTCCCTTAACTCCATCAGACCTCTTACGACCATTAAGACTCTGCGAGA 1248
 Db GCTTTACCTCTGACTCCCTTAACTCCATCAGACCTCTTACGACCATTAAGACTCTGCGAGA 1107
 QY ACTGAGAGCC 1259
 Db ACTGAGAGGCC 1118

RESULT 3
 LOCUS BX420015
 DEFINITION BX420015 Homo sapiens FETAL BRAIN Homo sapiens cdna clone
 CS0DF023YA09 5-PRIME, mRNA sequence.
 ACCESSION BX420015
 VERSION BX420015.2 GI:46929710
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 1062)
 Li, W. B., Gruber, C., Jessee, J., and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On May 13, 2003 this sequence version replaced gi:30646738.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
 was not normalized. Library was constructed by Life Technologies, a
 division of Invitrogen.
 This sequence belongs to sequence cluster 6147.r
 For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna?s=CS0DF023YA09&c=6147.r.

FEATURES
 Location/Qualifiers
 1..1062
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DF023YA09"
 /tissue_type="FETAL BRAIN"
 /dev_stage="fetal"
 /clone_lib="Homo sapiens FETAL BRAIN"
 /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
 was primed with a NotI-oligo(dT) primer. Five prime end
 enriched, double-strand cDNA was digested with Not I and
 cloned into the Not I and EcoRV sites of the pCMVSPORT 6
 vector. Library was not normalized."

ORIGIN

Query Match 47.8%; Score 600; DB 5; Length 1062;
 Best Local Similarity 99.7%; Pred. No. 1.3e-302;
 Matches 700; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 349 AGCTTTGGAAGCGCCAGAGAGTCCCTGCTTCAATAGATTGTTTCCCTCGCT 408
 Db AGCTTTGGAAGCGCCAGAGAGTCCCTGCTTCAATAGATTGTTTCCCTCGCT 266
 QY 409 TCTCTGCTGTTATCTACTGGGTGAGTGTCTGCTTCCCTGTGTGTGGAGTGCCTCG 468
 Db TCTCTGCTGTTATCTACTGGGTGAGTGTCTGCTTCCCTGTGTGTGGAGTGCCTCG 326
 QY 469 GAGACGAGGCGGTGCGAGGCAACCCCATGAAGCTGCGTGCATCTCTGATGAAGAGA 528
 Db GAGACGAGGCGGTGCGAGGCAACCCCATGAAGCTGCGTGCATCTCTGATGAAGAGA 386
 QY 529 GAGGAGGTGGAGGCGACACGCTGGTGAATGGTTCTACAGGCCGAGGCGGTAAAGAT 588
 Db GAGGAGGTGGAGGCGACACGCTGGTGAATGGTTCTACAGGCCGAGGCGGTAAAGAT 446
 QY 589 TTCTCTTATTTACGAGTATCGGAATGGCCACAGGAGGTGAGAGCCCTTTTCAGGGGCGC 648
 Db TTCTCTTATTTACGAGTATCGGAATGGCCACAGGAGGTGAGAGCCCTTTTCAGGGGCGC 506
 QY 649 CTGACGTGGAATGCGCAGAGGACCTGACAGGACGTGTCCATCACTGTGCTCAACGTCACT 708
 Db CTGACGTGGAATGCGCAGAGGACCTGACAGGACGTGTCCATCACTGTGCTCAACGTCACT 566
 QY 709 CTGAACGACTCTGSCCTCTACACTGCAATGTGTCCCGGAGTTTGAGTTTGAGGCGCAT 768
 Db CTGAACGACTCTGSCCTCTACACTGCAATGTGTCCCGGAGTTTGAGTTTGAGGCGCAT 626
 QY 769 CGGCGCTTTGTGAAGACGACGCGGCTGATCCCCCTAAGAGTCAACGAGGAGGCTGGAGAG 828
 Db CGGCGCTTTGTGAAGACGACGCGGCTGATCCCCCTAAGAGTCAACGAGGAGGCTGGAGAG 686
 QY 829 GACTTCACTCTGTGGTCTCAGAAATCATGATGATGATGATGATGATGATGATGATGATGAT 888
 Db GACTTCACTCTGTGGTCTCAGAAATCATGATGATGATGATGATGATGATGATGATGATGAT 746

cloning). Average insert size 1.7 kb, insert size range 0.7-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 017. Note: this is a NIH_MGC Library."

```

ORIGIN
Query Match      22.8%; Score 288; DB 5; Length 1069;
Best Local Similarity 100.0%; Pred. No. 4.9e-139;
Matches 288; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 349 AGCCTTGGAGCGCCAGCCAGAGATGCTTCAATAGATTGTTCCCTGGCT 408
DB 355 AGCCTTGGAGCGCCAGCCAGAGATGCTTCAATAGATTGTTCCCTGGCT 414
QY 409 TCTCTGCTTATCTACTAGGTCAGTGTCTGCTTCCCTGTGTGTGGAAGTGCCTCG 468
DB 415 TCTCTGCTTATCTACTAGGTCAGTGTCTGCTTCCCTGTGTGTGGAAGTGCCTCG 474
QY 469 GAGACGAGCGCGTGCAGGGCAACCCATGAAGTGCCTGCATCTCTGCATGAAGAGA 528
DB 475 GAGACGAGCGCGTGCAGGGCAACCCATGAAGTGCCTGCATCTCTGCATGAAGAGA 534
QY 529 GAGGAGTGGAGCCACACGCTGGTGGATGGTCTACAGCCGAGGCGGTAAAGAT 588
DB 535 GAGGAGTGGAGCCACACGCTGGTGGATGGTCTACAGCCGAGGCGGTAAAGAT 594
QY 589 TTCTTTATTTAGAGTATCGGAATGCCACAGGAGTGGAGAGCCCC 636
DB 595 TTCTTTATTTAGAGTATCGGAATGCCACAGGAGTGGAGAGCCCC 642

```

```

RESULT 12
LOCUS      AL5341136
DEFINITION AL5341136 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
            CS0DF005Y102 5-PRIME, mRNA sequence.
ACCESSION  AL5341136
VERSION     AL5341136.3 GI:45709952
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
            Full-length cDNA libraries and normalization
            Unpublished (2001)
            On Feb 13, 2001 this sequence version replaced gi:30539643.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with NotI and cloned
            into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library
            was not normalized. Library was constructed by Life Technologies, a
            division of Invitrogen.
            This sequence belongs to sequence cluster 6147.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?8=CS0DF005BB01QP1&c=6147.r.
            Location/Qualifiers
            1. .1078
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="CS0DF005Y102"
            /tissue_type="FETAL BRAIN"
            /dev_stage="fetal"
            /clone_lib="Homo sapiens FETAL BRAIN"
            /notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
            was primed with a NotI-oligo(dT) primer. Five prime end
            enriched, double-strand cDNA was digested with NotI and

```

```

FEATURES
source
1. .1078
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF005Y102"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with NotI and

```

cloned into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

```

ORIGIN
Query Match      21.3%; Score 268; DB 1; Length 1078;
Best Local Similarity 99.7%; Pred. No. 1.5e-128;
Matches 388; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 134 GGCTGATCGGCTCCCTCGAACTGGGGAGGTCCAGTGGGTGCTTTAGGSCCCCAAGCCCC 193
DB 521 GGCTGATCGGCTCCCTCGAACTGGGGAGGTCCAGTGGGTGCTTTAGGSCCCCAAGCCCC 580
QY 194 GACCGGCTCCAAAAGTCCCAAGGGCTCCCAAGCAACCGTGTCTCGGCCCTTCTTCGG 253
DB 581 CACCGGCTCCAAAAGTCCCAAGGGCTCCCAAGCAACCGTGTCTCGGCCCTTCTTCGG 640
QY 254 TCAGAAAGTTCGGCCCTCTGGGGCAGTTCTGTCCCAAGGGTTTCTCGAAAGAAATCTGAGA 313
DB 641 TCAGAAAGTTCGGCCCTCTGGGGCAGTTCTGTCCCAAGGGTTTCTCGAAAGAAATCTGAGA 700
QY 314 GGGCGCAGTCTCTGACCGAGGGAATCTCTGTGTAGCTTCGAAAGCCGCCAGCCCCAGA 373
DB 701 GGGCGCA-TCTCTTGACCGAGGGAATCTCTGTGTAGCTTCGAAAGCCGCCAGCCCCAGA 759
QY 374 AGATCGCTGCCTTCAATAGATTGTTCCCTGGCTTCTCTGCTTACTACTGGGTCA 433
DB 760 AGATCGCTGCCTTCAATAGATTGTTCCCTGGCTTCTCTGCTTACTACTGGGTCA 819
QY 434 GTGTCTGTCTTCCCTGTGTGTGGAAGTCCCTCGAGACGAGGCGCTGCAGGGCAACC 493
DB 820 GTGTCTGTCTTCCCTGTGTGTGGAAGTCCCTCGAGACGAGGCGCTGCAGGGCAACC 879
QY 494 CCATGAAGTGGCTGCATCTCTCGATG 522
DB 880 CCATGAAGTGGCTGCATCTCTCGATG 908

```

```

RESULT 13
LOCUS      D44825
DEFINITION HUMSUPY279 Human brain cDNA Homo sapiens cDNA clone NF220-K, mRNA
            sequence.
ACCESSION  D44825
VERSION     D44825.1 GI:1572300
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            Hadano, S., Ishida, Y., Tomiyasu, H., Yamamoto, K., Bates, G.P. and
            Ikeda, J.
            Transcript map of the human chromosome 4p16.3 consisting of 627
            cDNA clones derived from 1 Mb of the Huntington's disease locus
            DNA Res. 3 (4), 239-255 (1996)
            JOURNAL 97101646
            PUBMED 8946164
            COMMENT
            Contact: Shinji Hadano
            Japan Science and Technology Corporation, NeuroGenes Project, ICORP
            Univ. of Tokai School of Med.
            Bohseidai, Isehara, Kanagawa 259-1193, Japan
            Tel: 81-463-91-5095
            Fax: 81-463-91-4993
            Email: shinji@ng.med.u-tokai.ac.jp.
            Location/Qualifiers
            1. .443
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="NF220-K"
            /tissue_type="brain"
            /clone_lib="Human brain cDNA"

FEATURES
source
1. .443
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NF220-K"
/tissue_type="brain"
/clone_lib="Human brain cDNA"

ORIGIN

```

```
Query Match      13.1%; Score 165; DB 7; Length 443;
Best Local Similarity 99.5%; Pred. No. 1.5e-74;
Matches 215; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1044 AGGTGGCTGAACACCTGAGGAGTGGACATCCCATGTTTCAGCAATGCAATGCGATCAG 1103
Db 33 AGGTGGCTGAACACCTGAGGAGTGGACATCCCATGTTTCAGCAATGCAATGCGATCAG 92

QY 1104 GAGGGCCCCCAAGGGCCCCCATCGCTTCCCTTCATGATCCATGTTCTGTTCAATTCATT 1163
Db 93 GAGGGCCCCCAAGGGCCCCCATCGCTTCCCTTCATGATCCATGTTCTGTTCAATTCATT 152

QY 1164 CATCTATACATCCACATGCGCTCTGAGCTTTTCACCTCTGACTCCCTAACTCCATCAGACCT 1223
Db 153 CATCTATACATCCACATGCTCTGAGCTTTTCACCTCTGACTCCCTAACTCCATCAGACCT 212

QY 1224 CTAGGCACCATAGACTCTGCCAGAACTGAGAGCC 1259
Db 213 CTAGGCACCATAGACTCTGCCAGAACTGAGAGCC 248

RESULT 14
AY419146
LOCUS      584 bp DNA linear GSS 12-DEC-2003
DEFINITION Pan troglodytes HCM6793 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION  AY419146
VERSION     AY419146.1 GI:39775106
KEYWORDS   GSS.
SOURCE     Pan troglodytes (chimpanzee)
ORGANISM   Pan troglodytes
REFERENCE   1 (bases 1 to 584)
AUTHORS    Clark,A.G., Gnanawski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE      Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL    Science 302 (5652), 1960-1963 (2003)
PUBMED    14671302
REFERENCE   2 (bases 1 to 584)
AUTHORS    Clark,A.G., Gnanawski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE      Direct Submission
JOURNAL    Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT    These sequences were made by sequencing genomic exons and ordering them based on alignment.
FEATURES   source
            1..584
            /organism="Pan troglodytes"
            /mol_type="genomic DNA"
            /db_xref="taxon:9598"
            <1..584
            /locus_tag="HCM6793"

gene
ORIGIN
Query Match      12.5%; Score 158; DB 9; Length 584;
Best Local Similarity 100.0%; Pred. No. 6.9e-71;
Matches 158; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 376 ATGCCTGCCTCAATAGATTGTTTCCCTCGCTTCTCTCGTGTCTATCTACTGGGTCAGT 435
Db 1 ATGCCTGCCTCAATAGATTGTTTCCCTCGCTTCTCTCGTGTCTATCTACTGGGTCAGT 60

QY 436 GTCTGCTTCCGTGTGTGTGGAGTGCCTTCGAGACGGAGCCGTCGAGGGCAACCC 495
Db 61 GTCTGCTTCCGTGTGTGTGGAGTGCCTTCGAGACGGAGCCGTCGAGGGCAACCC 120
```

```
QY 496 ATGAAGCTGGCTGATCTCTCTGCATGAAGAGAGGA 533
Db 121 ATGAAGCTGGCTGATCTCTCTGCATGAAGAGAGGA 158

RESULT 15
BX280396
LOCUS      509 bp mRNA linear EST 04-MAR-2003
DEFINITION BX280396 NIH_MGC_121 Homo sapiens cDNA clone IMAGE9998F0412832; IMAGE:5769987, mRNA sequence.
ACCESSION  BX280396
VERSION     BX280396.1 GI:28612393
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE   1 (bases 1 to 509)
AUTHORS    Ebert,L., Heil,O., Hennig,S., Neubert,P., Partsch,E., Peters,M., Radelof,U., Schneider,D. and Korn,B.
TITLE      Human Unigeneset - RZPD3
JOURNAL    Unpublished (2003)
COMMENT    Contact: Ina Rolfs
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
            RZPD: IMAGE998F0412832.
            RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
            Human Unigeneset - RZPD3 (RZPDLIB No.972)
            http://www.rzpd.de/CloneCards/cgi-bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rolfs
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 101
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available royalty-free from RZPD;
            contact RZPD (clone@rzpd.de) for further information. Seq primer:
            M13u, Primer sequence: CGTGTAAACCGACGCCAGT.

FEATURES   Location/Qualifiers
            1..509
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE998F0412832; IMAGE:5769987"
            /lab_host="DH10B"
            /clone_lib="NIH_MGC_121"
            /note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 3 fetal brains, female age 20 weeks, female age 24 weeks, and male age 26 weeks. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 0.7-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 017. Note: this is a NIH_MGC Library."

ORIGIN
Query Match      12.2%; Score 154; DB 5; Length 509;
Best Local Similarity 100.0%; Pred. No. 8.6e-69;
Matches 154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 349 AGCCTTGGAGCCGCCAGCCAGAGATGCCCTTCAATAGATTGTTTCCCTGGCT 408
Db 356 AGCCTTGGAGCCGCCAGCCAGAGATGCCCTTCAATAGATTGTTTCCCTGGCT 415

QY 409 TCCTCTCGTGTATCTACTAGGTCAGTGTCTGCTTCCCTGTGTGTGGAAGTCCCTCG 468
Db 416 TCCTCTCGTGTATCTACTAGGTCAGTGTCTGCTTCCCTGTGTGTGGAAGTCCCTCG 475

QY 469 GAGACGGAGCCGCTGAGGGCAACCCCATGAAGC 502
```

Db 476 GAGACGGGGCGTGCAGGGGCAACCCCATGAAGC 509

Search completed: April 7, 2005, 08:50:28
Job time : 3950 secs